

STIC-Biotech/ChemLib

193461

From: Chan, Christina
Sent: Wednesday, June 21, 2006 8:39 AM
To: Swope, Sheridan; STIC-Biotech/ChemLib
Subject: RE: 10/825,632

~~Please rush. Thanks Chris~~

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

WRFE

-----Original Message-----

From: Swope, Sheridan
Sent: Tuesday, June 20, 2006 7:03 PM
To: Chan, Christina
Subject: 10/825,632

Chris, May I have this rushed for an allowance?

Please search:

SID 2, oligo search, against the GenEmbl and EST databases only

Collect the top 150 hits from each search.

Sheridan Swope, Ph.D.
Primary Patent Examiner
AU 1656/Recombinant Enzymes
571-272-0943 (voice)
E02D19 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)
Helping applicants get good patents.

79280

BEST AVAILABLE COPY

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 193461

TO: Sheridan Swope
Location: REM-2B71&3C70
Art Unit: 1656
Thursday, June 22, 2006
Case Serial Number: 10/825632

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1a69
Phone: 571-272-2518

BOB
barbara.obryen@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

Thu Jun 22 09:04:26 2006

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 22, 2006, 03:28:17 ; Search time 14067 Seconds
(without alignments)
12402.667 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 3120
Sequence: 1 aadgttaagctcgagg.....agaattactaaaaaaaaa 3120

Scoring table: OLIGO NUC
Gapop '60.0 , Gapext 60.0

Searched: 48236798 seqs, 27959665780 residues

Word size : 1
Total number of hits satisfying chosen parameters: 96473154

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database :

- 1: gb_est1:*
- 2: gb_est3:*
- 3: gb_est4:*
- 4: gb_est5:*
- 5: gb_est6:*
- 6: gb_hic:*
- 7: gb_est2:*
- 8: gb_est7:*
- 9: gb_est8:*
- 10: gb_est9:*
- 11: gb_ges1:*
- 12: gb_ges2:*
- 13: gb_ges3:*
- 14: gb_ges4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2767	88.7	4535	6	BC040203
2	1664	53.3	2292	6	CR609512 full-length
3	966	31.0	2649	14	AY411615 Homo sapi
4	782	25.1	1292	6	AF175225 Homo sapi
5	769	24.6	1042	3	BM557438 AGENCOURT
6	755	24.2	884	8	CR985650 CR985650
7	726	23.3	910	2	BG479035 AGENCOURT
8	725	23.2	968	3	BQ671635 AGENCOURT
9	724	23.2	910	3	BQ675260 AGENCOURT
10	723	23.1	1265	6	AF176779 Homo sapi
11	722	22.7	7351	1	AL043338 Homo sapi
12	708	22.4	753	1	AL040398 Homo sapi
13	700	22.1	835	8	CR998849 Homo sapi
14	690	22.1	830	9	DA571902 Homo sapi
15	689	22.0	742	8	CK000768 Homo sapi
16	687	22.0	914	4	EX390898 Homo sapi
17	686	22.0	724	5	CK000768 Homo sapi
18	685	22.0	724	5	CK000768 Homo sapi
19	682	21.9	820	8	CR994557 Homo sapi

20	676	21.7	864	1	AL542617
21	674	21.6	2649	14	AY411616
22	672	21.5	1002	7	BE733691
23	671	21.5	671	9	CX787143
24	663	21.2	804	3	BU687687
25	660	21.2	993	2	BG259714
26	652	20.9	910	4	BX372276
27	649	20.8	657	9	DA415588
28	647	20.7	789	2	BG709118
29	644	20.6	669	2	CG390103
30	641	20.5	881	8	CR990494
31	639	20.5	708	3	BU622228
32	637	20.4	722	3	BU631054
33	627	20.1	853	4	BI223892
34	626	20.1	815	4	CB990233
35	626	20.1	826	2	BI553405
36	623	20.0	695	2	BG715866
37	614	19.7	634	1	AI819365
38	607	19.5	957	3	BQ675006
39	598	19.2	612	8	CV024163
40	596	19.1	596	4	CB153590
41	596	19.1	596	4	DA190607
42	587	18.8	587	9	DA392243
43	586	18.8	586	9	DA441156
44	585	18.8	585	9	DA157274
45	583	18.7	583	9	DA159782
46	582	18.7	582	9	DA488030
47	581	18.6	581	9	DA390016
48	580	18.6	580	9	DA645431
49	580	18.6	580	9	DA168729
50	580	18.6	580	9	DA1084090
51	578	18.5	593	9	DA156844
52	578	18.5	629	9	DA140770
53	576	18.5	576	9	DA301389
54	576	18.5	576	9	DA690898
55	576	18.5	577	9	DA478474
56	576	18.5	627	2	BG752328
57	575	18.4	575	9	DA163570
58	573	18.4	573	9	DA769628
59	573	18.4	574	9	DA931907
60	571	18.3	573	9	DA858700
61	562	18.0	760	1	AI917735
62	561	18.0	563	8	DA860616
63	561	18.0	678	8	DA27235
64	558	17.9	673	7	AW303607
65	557	17.9	557	9	DA997805
66	556	17.8	558	9	DA756453
67	555	17.8	788	2	BG163397
68	553	17.7	553	9	DB074943
69	552	17.7	566	9	DA268082
70	548	17.6	553	8	DA427230
71	548	17.6	612	8	DA427227
72	547	17.5	1278	6	AF173382
73	546	17.5	546	9	DA052731
74	543	17.4	897	3	BU191638
75	541	17.3	541	9	DA809185
76	539	17.3	556	9	DB339434
77	539	17.3	540	9	DA437440
78	539	17.3	542	9	DA396556
79	539	17.3	569	9	DB199007
80	539	17.3	551	9	DA446578
81	538	17.2	597	3	BU784521
82	534	17.1	534	9	DA515446
83	534	17.1	585	9	DA927284
84	527	16.9	578	9	DA773829
85	526	16.9	674	7	BE888665
86	525	16.8	565	9	DA694456
87	525	16.8	717	9	DA828964
88	523	16.8	523	9	DR002989
89	522	16.7	546	9	DA146190
90	521	16.7	523	9	DA990713
91	521	16.6	542	8	CN427232
92	518	16.6	639	7	AW958665

93	-517	16.6	517	9	DA412358	DA412358	DA412358
94	515	16.5	566	9	DA817784	DA817784	DA817784
95	515	16.5	568	7	BE388695	601286345	BE388695
96	513	16.4	564	9	DA766261	DA766261	DA766261
97	511	16.4	563	9	DA033029	DA033029	DA033029
98	510	16.3	559	9	DB020491	DB020491	DB020491
99	510	16.3	575	9	DB242468	DB242468	DB242468
100	509	16.3	560	9	DA523081	DA523081	DA523081
101	508	16.3	611	9	DA437193	DA437193	DA437193
102	508	16.3	547	9	DB041170	DB041170	DB041170
103	508	16.3	559	9	DB370407	DB370407	DB370407
104	508	16.3	559	9	DA956211	DA956211	DA956211
105	502	16.1	553	9	DA450229	DA450229	DA450229
106	501	16.1	531	9	DA831627	DA831627	DA831627
107	500	16.0	642	2	BI546487	603198793	BI546487
108	496	15.9	704	8	BI442726	170004553	BI442726
109	-495	15.8	500	1	AA496257	aa24b08.8	AA496257
110	-493	15.8	500	1	AA496257	aa24b08.8	AA496257
111	-492	15.8	617	7	BE789604	601481552	BE789604
112	481	15.5	639	2	BG165205	602344074	BG165205
113	481	15.4	553	9	DA522231	DA522231	DA522231
114	481	15.4	536	2	BI710784	id90808.Y	BI710784
115	479	15.4	855	2	BI084885	60289453	BI084885
116	478	15.3	661	7	BE799580	601532450	BE799580
117	476	15.3	577	1	AI243857	gh89f09.X	AI243857
118	472	15.1	575	9	DA798581	DA798581	DA798581
119	468	15.0	521	9	DA189698	DA189698	DA189698
120	-457	14.6	457	1	AA306656	EST177662	AA306656
121	455	14.6	576	9	DB156905	DB156905	DB156905
122	452	14.5	587	9	DA607747	DA607747	DA607747
123	449	14.4	570	9	DA145947	DA145947	DA145947
124	449	14.4	570	9	DB235712	DB235712	DB235712
125	447	14.3	568	9	DA082837	DA082837	DA082837
126	447	14.3	570	9	DB161177	DB161177	DB161177
127	-445	14.3	469	1	AA465309	oc93h12.s	AA465309
128	437	14.0	437	1	AA465309	oc93h12.s	AA465309
129	437	14.0	558	9	DA455896	DA455896	DA455896
130	434	13.9	726	8	CA427231	170004554	CA427231
131	433	13.9	494	1	AA278626	DB358826	AA278626
132	430	13.8	514	9	DA358826	DB358826	DA358826
133	430	13.8	574	9	DA456165	DA456165	DA456165
134	424	13.6	548	9	DA649369	DA649369	DA649369
135	424	13.6	592	9	DA392311	DA392311	DA392311
136	418	13.4	857	2	BI462092	603205272	BI462092
137	416	13.3	474	5	CF137069	UI-HF-BN0	CF137069
138	413	13.2	570	9	DA297112	DA297112	DA297112
139	408	13.1	530	9	DA551227	DA551227	DA551227
140	408	13.1	569	9	DA436183	DA436183	DA436183
141	404	12.9	590	9	DA946649	DA946649	DA946649
142	404	12.9	835	2	BI548490	603191203	BI548490
143	399	12.8	399	1	AA769522	nz35c08.s	AA769522
144	398	12.8	598	9	DB204093	DB204093	DB204093
145	395	12.7	738	8	CA427234	170006000	CA427234
146	393	12.6	493	1	AA278625	z878g11.X	AA278625
147	392	12.6	392	1	AA312210	EST182961	AA312210
148	391	12.5	552	9	DB058900	DB058900	DB058900
149	388	12.4	798	2	BG719202	602690257	BG719202
150	384	12.3	555	9	DA184650	DA184650	DA184650

ALIGNMENTS

RESULT 1	BC040203	Homo sapiens	4535 bp	mRNA	linear	HTC	19-JAN-2006
LOCUS	BC040203	Homo sapiens	dipeptidyl-peptidase 8, mRNA (cDNA clone IMAGE:4824813), with apparent retained intron.				
DEFINITION	BC040203	Homo sapiens	dipeptidyl-peptidase 8, mRNA (cDNA clone IMAGE:4824813), with apparent retained intron.				
ACCESSION	BC040203	Homo sapiens	dipeptidyl-peptidase 8, mRNA (cDNA clone IMAGE:4824813), with apparent retained intron.				
VERSION	BC040203.1	Homo sapiens	dipeptidyl-peptidase 8, mRNA (cDNA clone IMAGE:4824813), with apparent retained intron.				
KEYWORDS	HTC.	Homo sapiens	dipeptidyl-peptidase 8, mRNA (cDNA clone IMAGE:4824813), with apparent retained intron.				
SOURCE	BC040203	Homo sapiens	dipeptidyl-peptidase 8, mRNA (cDNA clone IMAGE:4824813), with apparent retained intron.				
ORGANISM	BC040203	Homo sapiens	dipeptidyl-peptidase 8, mRNA (cDNA clone IMAGE:4824813), with apparent retained intron.				

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.	1 (bases 1 to 4535)
REFERENCE	
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, A., Whiting, M., Madan, A., Rodriguez, S., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butcherfield, Y.S., Krzywicki, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
CONSRMTM	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
TITLE	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL	12477932
REFERENCE	2 (bases 1 to 4535)
AUTHORS	NIH MGC Project
CONSRMTM	Direct Submission
TITLE	Submitted (27-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
JOURNAL	NIH-MGC Project URL: http://mgc.nci.nih.gov
REMARK	Contact: MGC help desk
COMMENT	Email: cgapb-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amg@bcm.tmc.edu Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowals, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAP Plate: 71 Row: K Column: 23 This clone has the following problem: retained intron.
Source	Location/Qualifiers 1..4535 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:4824813" /cissue_type="Testis" /clone_lib="NIH_MGC_97" /lab_host="DH10B" /note="Vector: pBluescriptR"
ORIGIN	
Query Match	88.7%; Score 2767; DB 6; Length 4535;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 2917; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
Qy	201 ACGAAATGCAACATGGCAGCAGCAATGGAAACAGAACAGACTGGGTGGATATTGA 260
Db	1616 ACGAAATGCAACATGGCAGCAGCAATGGAAACAGAACAGACTGGGTGGATATTGA 1675

Qy	261	AATCGGCGACTGTGAGGAGATATTGAATCACAGATCGGCTAAATTTGGAGCGCTTTTA	320	Db	2756	CACTCTACTAGATCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGGAATTATT	2815
Db	1676	AATCGGCGACTGTGAGGAGATATTGAATCACAGATCGGCTAAATTTGGAGCGCTTTTA	1735	Qy	1401	TATCCCACTAGAGATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCACTGCTGATTC	1460
Qy	321	TGTTGAGCGGTATTTCTGAGAGTCACTTAAAGGCTGCTTGCAGATACACGAAATATCA	380	Db	2816	TATCCCACTAGAGATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCACTGCTGATTC	2875
Db	1736	TGTTGAGCGGTATTTCTGAGAGTCACTTAAAGGCTGCTTGCAGATACACGAAATATCA	1795	Qy	1461	TGTGAGCGCACATAATTTCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGAT	1520
Qy	381	TGGCTACATGATGGCTAAGGACACACATGATTTCAATGTTTGTGAGAGGATGATCCAGA	440	Db	2876	TGTGAGCGCACATAATTTCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGAT	2935
Db	1796	TGGCTACATGATGGCTAAGGACACACATGATTTCAATGTTTGTGAGAGGATGATCCAGA	1855	Qy	1521	CTTTTCATGTTTTTCCCAAGTCAAGAGAGAAATTTGAGTTTATTTTTTCCCTCTCAATG	1580
Qy	441	TGGACCTCATTCAGACAGATCTATTACCTTCCATGTCGATGTCGATGAGAACAGAAATAC	500	Db	2936	CTTTTCATGTTTTTCCCAAGTCAAGAGAGAAATTTGAGTTTATTTTTTCCCTCTCAATG	2995
Db	1856	TGGACCTCATTCAGACAGATCTATTACCTTCCATGTCGATGTCGATGAGAACAGAAATAC	1915	Qy	1581	CAAAAAGGTTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATATA	1640
Qy	501	ACTGTTTTTATCTGAAATTCCTCAAAATCTATCAATAGAGCAGAGCTTTAATGCTCTTG	560	Db	2996	CAAAAAGGTTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATATA	3055
Db	1916	ACTGTTTTTATCTGAAATTCCTCAAAATCTATCAATAGAGCAGAGCTTTAATGCTCTTG	1975	Qy	1641	ACGATCCAGTGGTGGCTCGCTGCCCAAGTGATTTCAAGTGTCTATCAAGAGGAGAT	1700
Qy	561	GAAAGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAAGA	620	Db	3056	ACGATCCAGTGGTGGCTCGCTGCCCAAGTGATTTCAAGTGTCTATCAAGAGGAGAT	3115
Db	1976	GAAAGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAAGA	2035	Qy	1701	AGCAATTTACAGTGGTGAATGGAAGTCTTTGGCCGCGCATGATCTAATATCCAAAGTTGA	1760
Qy	621	AGAACTATTAAAGAGAAAGAAACGATTGGAAACAGTCGGAATTTGCTTCTTACGATTATCA	680	Db	3116	AGCAATTTACAGTGGTGAATGGAAGTCTTTGGCCGCGCATGATCTAATATCCAAAGTTGA	3175
Db	2036	AGAACTATTAAAGAGAAAGAAACGATTGGAAACAGTCGGAATTTGCTTCTTACGATTATCA	2095	Qy	1761	TGAAGTCAGAGAGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTTGAAGCATCACCT	1820
Qy	681	CCAGAGAGTGGAAATTTCTGTTTCAAGCGGTGTGGAATTTATCAGTAAAGATGG	740	Db	3176	TGAAGTCAGAGAGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTTGAAGCATCACCT	3235
Db	2096	CCAGAGAGTGGAAATTTCTGTTTCAAGCGGTGTGGAATTTATCAGTAAAGATGG	2155	Qy	1821	GTACGTAGTCAGATTACGTAAATCTCGAGAGGTGACAGAGCTGACTGACCGTGGTACTC	1880
Qy	741	AGGGCCACAAGGATTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTTC	800	Db	3236	GTACGTAGTCAGATTACGTAAATCTCGAGAGGTGACAGAGCTGACTGACCGTGGTACTC	3295
Db	2156	AGGGCCACAAGGATTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTTC	2215	Qy	1881	ACATTTCTGCTGCATCAGTCACTGCTGACTGCTTTTAACTAACTAGTATAGTAAACAGAA	1940
Qy	801	CAACATACCGATGGATCCAAATTTATGCCCTGCTGATCCAGATCGAATGCTTTTATACA	860	Db	3296	ACATTTCTGCTGCATCAGTCACTGCTGACTGCTTTTAACTAACTAGTATAGTAAACAGAA	3355
Db	2216	CAACATACCGATGGATCCAAATTTATGCCCTGCTGATCCAGATCGAATGCTTTTATACA	2275	Qy	1941	GAAATCCACATGCTGTGCTCCCTTTTCAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAA	2000
Qy	861	TAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAGAGGACTCACTTATGT	920	Db	3356	GAAATCCACATGCTGTGCTCCCTTTTCAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAA	3415
Db	2276	TGGCAACGATATTTGGATATCTAACATCGTAAACAGAGAGAGGACTCACTTATGT	2335	Qy	2001	AAACAAAGGAAATTTTGGGCCACCATTTTGGGATCAGCAGGTCTCTCTGACTATCTCC	2060
Qy	921	GCACATGAGCTAGCCAAACATCGAAGAGATGCCAGATCAGCTGGAGTGCCTACCTTTGT	980	Db	3416	AAACAAAGGAAATTTTGGGCCACCATTTTGGGATCAGCAGGTCTCTCTGACTATCTCC	3475
Db	2336	GCACATGAGCTAGCCAAACATCGAAGAGATGCCAGATCAGCTGGAGTGCCTACCTTTGT	2395	Qy	2061	TCACAGAAATTTTCTCTTTTGGAAAGTACTTCTGGAATTTACATTTGATGGGATGCTCTACAA	2120
Qy	981	TCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGTGTCCTCAAAAGCTGAAACAC	1040	Db	3476	TCACAGAAATTTTCTCTTTTGGAAAGTACTTCTGGAATTTACATTTGATGGGATGCTCTACAA	3535
Db	2396	TCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGTGTCCTCAAAAGCTGAAACAC	2455	Qy	2121	GCCTCATGATCTACAGCCTCGGAAAGAAATATCTCTACTGCTGCTGCTCATATATGCTGCTCC	2180
Qy	1041	TCCAGTGGTGGTAAATTTCTAGAAATTTCTATACGAAATATGATGATCTGAGGTGGA	1100	Db	3536	GCCTCATGATCTACAGCCTCGGAAAGAAATATCTCTACTGCTGCTGCTCATATATGCTGCTCC	3595
Db	2456	TCCAGTGGTGGTAAATTTCTAGAAATTTCTATACGAAATATGATGATCTGAGGTGGA	2515	Qy	2181	TCAGGTGCAATGCTGGTGAATATCGGTTTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCT	2240
Qy	1101	AAATTTCTGATGATACCTCCCTATGTTGGAACAAAGAGGGGAGATCAATTCCTGATACC	1160	Db	3596	TCAGGTGCAATGCTGGTGAATATCGGTTTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCT	3655
Db	2516	AAATTTCTGATGATACCTCCCTATGTTGGAACAAAGAGGGGAGATCAATTCCTGATACC	2575	Qy	2241	AGCCTCTCTAGGTTATGCTGTTGATGATAGACAAACAGGGGATCTCTGTCAACGAGGCT	2300
Qy	1161	TAAACACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGTGTGATGATGATGATGATG	1220	Db	3656	AGCCTCTCTAGGTTATGCTGTTGATGATAGACAAACAGGGGATCTCTGTCAACGAGGCT	3715
Db	2576	TAAACACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGTGTGATGATGATGATGATG	2635	Qy	2301	TAAATTTGAAAGGCGCTTTTAAATATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGA	2360
Qy	1221	TGAAGAGAGGATCATAGATGTCATAGATAGGAACTAAATTCACCTTTTGAAGATTTCTATT	1280	Db	3716	TAAATTTGAAAGGCGCTTTTAAATATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGA	3775
Db	2636	TGAAGAGAGGATCATAGATGTCATAGATAGGAACTAAATTCACCTTTTGAAGATTTCTATT	2695	Qy	2361	AGGATCCCAATATCTAGCTTCTCGATATGATTTTCACTTTAGATTTAGTGTGGGATCCCA	2420
Qy	1281	TGAAGAGGTTGAATATTTGCGCAGAGCTGGATGCACTCTCGAGGGGAAATATGCTTGTGC	1340	Db	3776	AGGATCCCAATATCTAGCTTCTCGATATGATTTTCACTTTAGATTTAGTGTGGGATCCCA	3835
Db	2696	TGAAGAGGTTGAATATTTGCGCAGAGCTGGATGCACTCTCGAGGGGAAATATGCTTGTGC	2755	Qy	2421	CGGCTGTGCTTATCGGAGGATACCTCTCCCTGATGCAATTAATGAGAGGTCAGATATCTT	2480
Qy	1341	CATCTCTACTAGATCGCTCCAGACTCGGCTACAGATAGTGTGATCTCACCTGAATTTATT	1400				

```
Db 3836 CCGGTGGTCTCTATGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGTCAGATATCTT 3895
QY 2481 CAGGGTTCTATTGCTGGGGCCCGAGTCACCTCTGTGATCTTCTATGATACAGATACAC 2540
Db 3896 CAGGGTTCTATTGCTGGGGCCCGAGTCACCTCTGTGATCTTCTATGATACAGATACAC 3955
QY 2541 GGAAGCTTATATGGGTCACTGACAGAGTGAACAGGCTATTAATCTAGGATCTGTGGC 2600
Db 3956 GGAAGCTTATATGGGTCACTGACAGAGTGAACAGGCTATTAATCTAGGATCTGTGGC 4015
QY 2601 CATGCAAGCAGAAAAGTTCCCTCTGACCAATGCTTACCTCTTACATGATGTTTCCT 2660
Db 4016 CATGCAAGCAGAAAAGTTCCCTCTGACCAATGCTTACCTCTTACATGATGTTTCCT 4075
QY 2661 GGATGAGAATGTCCATTTTGCACATACAGGCTATTAATCTAGGATCTGTGGC 2720
Db 4076 GGATGAGAATGTCCATTTTGCACATACAGGCTATTAATCTAGGATCTGTGGC 4135
QY 2721 AAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACACAGCAATAGATTTCTGAATC 2780
Db 4136 AAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACACAGCAATAGATTTCTGAATC 4195
QY 2781 GGGAGACATTTAGTAACTGCTTTTGGACCTTCAAGAAAACCTTGGATCAGCTAT 2840
Db 4196 GGGAGACATTTAGTAACTGCTTTTGGACCTTCAAGAAAACCTTGGATCAGCTAT 4255
QY 2841 TGCTGCTTAAAGATGATTAATTTTGAACCTGTGTAGAACTCTCTGTATACACTGGCTA 2900
Db 4256 TGCTGCTTAAAGATGATTAATTTTGAACCTGTGTAGAACTCTCTGTATACACTGGCTA 4315
QY 2901 TTTAAACAAATGAGAGGTTTAAATCAACAGAAAAACAGAAATTTGATCATCATTTTGAAT 4375
Db 4316 TTTAAACAAATGAGAGGTTTAAATCAACAGAAAAACAGAAATTTGATCATCATTTTGAAT 4435
QY 2961 ACTGCGCATGTAACATCTACTCTCTGAAATAAATGTGGTCCATGACAGGGTCTACGGTT 3020
Db 4376 ACTGCGCATGTAACATCTACTCTCTGAAATAAATGTGGTCCATGACAGGGTCTACGGTT 4435
QY 3021 TGTGGTAAATCTTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTG 3080
Db 4436 TGTGGTAAATCTTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTG 4495
QY 3081 AGAGACCCAGCAATACCAATAGAAATTAATAAAAAA 3120
Db 4496 AGAGACCCAGCAATACCAATAGAAATTAATAAAAAA 4535
```

```
RESULT 2
CR609512
LOCUS
DEFINITION
    2292 bp mRNA linear HTC 21-JUL-2004
    full-length cDNA clone CS0DL005YD02 of B cells (Ramos cell line)
    Cot 25-normalized of Homo sapiens (human).
ACCESSION
    CR609512
VERSION
    CR609512.1 GI:50490319
KEYWORDS
    HTC; CNSLT cDNA.
SOURCE
    Homo sapiens
    ORGANISM
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
        Homidae; Homo;
        1 (bases 1 to 2292)
        Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
        Full-length cDNA libraries and normalization
        Unpublished
        Contact : Feng Liang Email : fliang@lifetech.com URL :
        http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
        Faraday Avenue
        2 (bases 1 to 2292)
        Genoscope.
        Direct Submission
        Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
        BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
        - Web : www.genoscope.cns.fr)
```

COMMENT

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life technologies, a division of Invitrogen.

FEATURES

source
Location/Qualifiers
1..2292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL005YD02"
/issue_type="B cells (Ramos cell line) Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

```
Query Match      53.3%; Score 1664; DB 6; Length 2292;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 472 GCCATGCTCGGTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTCCTCAAACTATC 531
Db 11 GCCATGCTCGGTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTCCTCAAACTATC 70
QY 532 AATAGAGCAGCAGTCTTAATGCTCTCTTGAAGCCCTCTTTGGATCTTTTTCAGGCAACA 591
Db 71 AATAGAGCAGCAGTCTTAATGCTCTCTTGAAGCCCTCTTTGGATCTTTTTCAGGCAACA 130
QY 592 CTGGACTATGGAATGTATTTCTCGAGAGAGAACTATTAAGAGAGAAAGAAACGCAATGGA 651
Db 131 CTGGACTATGGAATGTATTTCTCGAGAGAGAACTATTAAGAGAGAAAGAAACGCAATGGA 190
QY 652 ACAGTCGGAATTTGCTTCTTACGATTATCACCAAGGAGTGGAAACATTTCTGTTTCAAGCC 711
Db 191 ACAGTCGGAATTTGCTTCTTACGATTATCACCAAGGAGTGGAAACATTTCTGTTTCAAGCC 250
QY 712 GGTAGTGGAAATTTATCACGTAAGAGATGAGAGGCCACCAAGGATTTTACCCACACCTTTA 771
Db 251 GGTAGTGGAAATTTATCACGTAAGAGATGAGAGGCCACCAAGGATTTTACCCACACCTTTA 310
QY 772 AGGCCCAATCTAGTGGAAACTAGTCTCCCAACATACGAGTGGATCCCAAAATTTATGCC 831
Db 311 AGGCCCAATCTAGTGGAAACTAGTCTCCCAACATACGAGTGGATCCCAAAATTTATGCC 370
QY 832 GCTGATCCAGATCGGATGCTTTTATACATACGATATTTGGATATCTTAACATCGTA 891
Db 371 GCTGATCCAGATCGGATGCTTTTATACATACGATATTTGGATATCTTAACATCGTA 430
QY 892 ACCAGAGAGAGAGAGAGACTCACTTATGTCACATGAGCTAGCCAAACATGGAAGAGAT 951
Db 431 ACCAGAGAGAGAGAGAGACTCACTTATGTCACATGAGCTAGCCAAACATGGAAGAGAT 490
QY 952 GCCAGATCAGCTGGAGTCGCTACCTTTCTTCCAAAGAAATTTGATAGATTTCTGGC 1011
Db 491 GCCAGATCAGCTGGAGTCGCTACCTTTCTTCCAAAGAAATTTGATAGATTTCTGGC 550
QY 1012 TATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAAAATTTCTAGAAATCTA 1071
Db 551 TATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAAAATTTCTAGAAATCTA 610
QY 1072 TATGAAGAAATGATGAATCTGAGTGGAAATTTATCATGTTACATCCCTATGTTGGAA 1131
Db 611 TATGAAGAAATGATGAATCTGAGTGGAAATTTATCATGTTACATCCCTATGTTGGAA 670
QY 1132 ACAGAGGGGCGAGATTCATTCGGTTATCCTAAAAACAGGTACAGCAATCTTAAAGTCACT 1191
Db 671 ACAGAGGGGCGAGATTCATTCGGTTATCCTAAAAACAGGTACAGCAATCTTAAAGTCACT 730
QY 1192 TTTAGATGTCGAAATTAATGATGATGCTGAAAGGAGGATCATAGATGTCATAGATAAG 1251
Db 731 TTTAGATGTCGAAATTAATGATGATGCTGAAAGGAGGATCATAGATGTCATAGATAAG 790
QY 1252 GAACATAATCAACTTTTGAGATTCTATTTGAAGAGGAGTTGAATATATTTCCAGAGCTGA 1311
```

Db	791	GAACATAATCAACCTTTTGAGATTCTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGA	850	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
Qy	1312	TGGACCTCTGAGGGAATAATATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCCTA	1371	1 (bases 1 to 2649)
Db	851	TGAGCTCTGAGGGAATAATATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCCTA	910	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, P., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
Qy	1372	CAGATAGTGTGATCTCACTGAATATTTATCCAGTGAAGATCATGTTATGAAAGG	1431	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
Db	911	CAGATAGTGTGATCTCACTGAATATTTATCCAGTGAAGATCATGTTATGAAAGG	970	Science 302 (5652), 1960-1963 (2003)
Qy	1432	CAGAGACTCATTTAGTCACTGCTGATTTCTGTCAGCCACTAAATTTATCTATGAAGAA	1491	14671302
Db	971	CAGAGACTCATTTAGTCACTGCTGATTTCTGTCAGCCACTAAATTTATCTATGAAGAA	1030	2 (bases 1 to 2649)
Qy	1492	ACAGACATCTGGATAAATATCCATGACATCTTTTATGTTTCCCAAAGTCAAGAGAG	1551	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, P., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
Db	1031	ACAGACATCTGGATAAATATCCATGACATCTTTTATGTTTCCCAAAGTCAAGAGAG	1090	Direct Submission
Qy	1552	GAATTTGAGTTTATTTTGGCTCTGATGCAAAACAGGTTTCCGTCATTTATCAAAAT	1611	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
Db	1091	GAATTTGAGTTTATTTTGGCTCTGATGCAAAACAGGTTTCCGTCATTTATCAAAAT	1150	This sequence was made by sequencing genomic exons and ordering them based on alignment.
Qy	1612	ACATCTATTTAAAGGAAAGCAAAATATTAACGATCCAGTGGTGGCTGCTCCAAAGT	1671	Location/Qualifiers
Db	1151	ACATCTATTTAAAGGAAAGCAAAATATTAACGATCCAGTGGTGGCTGCTCCAAAGT	1210	1. .2649
Qy	1672	GATTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTT	1270	/organism="Homo sapiens"
Db	1211	GATTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTT	1270	/mol_type="genomic DNA"
Qy	1732	GGCCGGCATGATCTAATATCCAAAGTGTGATGAGTCCAGAGGCTCGTATATTTGAAGGC	1791	/db_xref="taxon:9606"
Db	1271	GGCCGGCATGATCTAATATCCAAAGTGTGATGAGTCCAGAGGCTCGTATATTTGAAGGC	1330	<1. .>2649
Qy	1792	ACCAAGATCCCTTTAGAGCATCACTGTAGTACGTTACGTTAAATCTCTGGAGAG	1851	/genes="DPP8"
Db	1331	ACCAAGATCCCTTTAGAGCATCACTGTAGTACGTTACGTTAAATCTCTGGAGAG	1390	/locus_tag="HCM4283"
Qy	1852	GTGACAAGCTGACTGACCGTGGCTACTCACTTCTGTCATGATCAGTCAGTGTGAC	1911	
Db	1391	GTGACAAGCTGACTGACCGTGGCTACTCACTTCTGTCATGATCAGTCAGTGTGAC	1450	
Qy	1912	TTCTTTAATAGTATAGTAAACGAGGAATCCACATGCTGTGTCCTTTTACAGCTA	1971	
Db	1451	TTCTTTAATAGTATAGTAAACGAGGAATCCACATGCTGTGTCCTTTTACAGCTA	1510	
Qy	1972	TCAAGTCTGGAAGATGACCAACTGTGCAAAACAAAGGAATTTTGGGCGACCATTTTGGAT	2031	
Db	1511	TCAAGTCTGGAAGATGACCAACTGTGCAAAACAAAGGAATTTTGGGCGACCATTTTGGAT	1570	
Qy	2032	TCAGCAGGTCCTCTTCTGACTATATCTCTCAGAAATTTCTTTTGAAGTACTACT	2091	
Db	1571	TCAGCAGGTCCTCTTCTGACTATATCTCTCAGAAATTTTCTTTTGAAGTACTACT	1630	
Qy	2092	GGATTTACATTTATGGATGCTCTTACAGCTCATGATCTACAGCTCGAAAGAAATAT	2151	
Db	1631	GGATTTACATTTATGGATGCTCTTACAGCTCATGATCTACAGCTCGAAAGAAATAT	1690	
Qy	2152	CCTACTGCTGTTTCATATATGTTGTTCTCTCAGGT	2186	
Db	1691	CCTACTGCTGTTTCATATATGTTGTTCTCTCAGGT	1725	
Qy	754	TTTACGCAACAACTTTTAAGGCCCAATCTAGTGGAACTAGTGTGTCCTCCCAACATACGGATG	813	
Qy	753	ACATTTCTGTTTCAAGCGGTAGTGAATTTATACGTTAAAGATGAGGGCCACAGGA	540	
Qy	481	ACATTTCTGTTTCAAGCGGTAGTGAATTTATACGTTAAAGATGAGGGCCACAGGA	540	
Qy	694	ACATTTCTGTTTCAAGCGGTAGTGAATTTATACGTTAAAGATGAGGGCCACAGGA	753	
Qy	421	GAAGAAACGCAATGGAACAGTCCGAATTTGCTTTACGATTATACCAAGGAAGTGA	480	
Qy	634	GAAGAAACGCAATGGAACAGTCCGAATTTGCTTTACGATTATACCAAGGAAGTGA	693	
Qy	574	GATCTTTTTCAGGCAACACTGGACTATGGAATGTTTCTCGAGAGGAAGTATTAAGA	633	
Qy	361	GATCTTTTTCAGGCAACACTGGACTATGGAATGTTTCTCGAGAGGAAGTATTAAGA	420	
Qy	514	GAAATTTCCCAAACTATCAATAGACGACGCTCTTAATGCTCTCTTGGAGGCTCTTTTG	360	
Qy	301	GAAATTTCCCAAACTATCAATAGACGACGCTCTTAATGCTCTCTTGGAGGCTCTTTTG	300	
Qy	241	GACAGAACTATTTACCTTCCCATGCTGCTGGTGAAGCAAGAAATACATGTTTATCT	240	
Qy	181	GCTAAGGCAACCATGATTTTCATGTTTGTGAAGGAATGATCCAGATGACCTCATTC	180	
Qy	394	GCTAAGGCAACCATGATTTTCATGTTTGTGAAGGAATGATCCAGATGACCTCATTC	393	
Qy	121	TCCTGGAGTACGCTTAAAGAGCTGCTTCCGATACCAAGAAATATCATGCTTACATG	120	
Qy	334	TCCTGGAGTACGCTTAAAGAGCTGCTTCCGATACCAAGAAATATCATGCTTACATG	333	
Qy	274	GAGGAGAAATTTGAATCAGAGGATCGGCTTAAATTTGGAGGCTTTTATGTTGAGCGGTAT	273	
Qy	61	GAGGAGAAATTTGAATCAGAGGATCGGCTTAAATTTGGAGGCTTTTATGTTGAGCGGTAT	60	
Qy	214	ATGCGACAGCAATGGAACAGAGCAAGCTGGTGTGTGAGATATTTGAACCTGCGAGCTGT	213	
Qy	1	ATGCGACAGCAATGGAACAGAGCAAGCTGGTGTGTGAGATATTTGAACCTGCGAGCTGT	1	
Qy	274	GAGGAGAAATTTGAATCAGAGGATCGGCTTAAATTTGGAGGCTTTTATGTTGAGCGGTAT	273	
Qy	61	GAGGAGAAATTTGAATCAGAGGATCGGCTTAAATTTGGAGGCTTTTATGTTGAGCGGTAT	60	
Qy	334	TCCTGGAGTACGCTTAAAGAGCTGCTTCCGATACCAAGAAATATCATGCTTACATG	333	
Qy	121	TCCTGGAGTACGCTTAAAGAGCTGCTTCCGATACCAAGAAATATCATGCTTACATG	120	
Qy	394	GCTAAGGCAACCATGATTTTCATGTTTGTGAAGGAATGATCCAGATGACCTCATTC	393	
Qy	181	GCTAAGGCAACCATGATTTTCATGTTTGTGAAGGAATGATCCAGATGACCTCATTC	180	
Qy	454	GACAGAACTATTTACCTTCCCATGCTGCTGGTGAAGCAAGAAATACATGTTTATCT	453	
Qy	241	GACAGAACTATTTACCTTCCCATGCTGCTGGTGAAGCAAGAAATACATGTTTATCT	240	
Qy	514	GAAATTTCCCAAACTATCAATAGACGACGCTCTTAATGCTCTCTTGGAGGCTCTTTTG	513	
Qy	301	GAAATTTCCCAAACTATCAATAGACGACGCTCTTAATGCTCTCTTGGAGGCTCTTTTG	300	
Qy	574	GATCTTTTTCAGGCAACACTGGACTATGGAATGTTTCTCGAGAGGAAGTATTAAGA	573	
Qy	361	GATCTTTTTCAGGCAACACTGGACTATGGAATGTTTCTCGAGAGGAAGTATTAAGA	360	
Qy	634	GAAGAAACGCAATGGAACAGTCCGAATTTGCTTTACGATTATACCAAGGAAGTGA	633	
Qy	421	GAAGAAACGCAATGGAACAGTCCGAATTTGCTTTACGATTATACCAAGGAAGTGA	420	
Qy	694	ACATTTCTGTTTCAAGCGGTAGTGAATTTATACGTTAAAGATGAGGGCCACAGGA	693	
Qy	481	ACATTTCTGTTTCAAGCGGTAGTGAATTTATACGTTAAAGATGAGGGCCACAGGA	480	
Qy	754	TTTACGCAACAACTTTTAAGGCCCAATCTAGTGGAACTAGTGTGTCCTCCCAACATACGGATG	753	

```

Db      541 TTTAGGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCTCCACATAGGATG 600
Qy      814 GATCAAAATATATCCCGCTGATCAGAGCTGGATTGCTTTTATACATAGCAAGATATT 873
Db      601 GATCAAAATATATCCCGCTGATCAGAGCTGGATTGCTTTTATACATAGCAAGATATT 660
Qy      874 TGGATATCTAACATCGTAACAGAGAAAGAGAGACTCACTTATGTGCAATGAGCTTA 933
Db      661 TGGATATCTAACATCGTAACAGAGAAAGAGAGACTCACTTATGTGCAATGAGCTTA 720
Qy      934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGCTTCCAAAGAA 993
Db      721 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGCTTCCAAAGAA 780
Qy      994 TTTGATAGATATTTCTGGCTATTGGTGTGTCCTCAAGCTGAAACCTCCAGTGGTGT 1053
Db      781 TTTGATAGATATTTCTGGCTATTGGTGTGTCCTCAAGCTGAAACCTCCAGTGGTGT 840
Qy      1054 AAAATCTTTAGAAATCTATATGAAGAAATGATGAATCTGAGTGGAAATTTATCATGTT 1113
Db      841 AAAATCTTTAGAAATCTATATGAAGAAATGATGAATCTGAGTGGAAATTTATCATGTT 900
Qy      1114 ACATCCCTTATGTTGGAACAGAGGAGGCGAGATTCATCCGTTATCCTTAAACAGGTACA 1173
Db      901 ACATCCCTTATGTTGGAACAGAGGAGGCGAGATTCATCCGTTATCCTTAAACAGGTACA 960
Qy      1174 GCAAAATCTTAAAGTCACTTTTAAGATGTCGAAATATGATGATGCTGAAGGAAGG 1230
Db      961 GCAAAATCTTAAAGTCACTTTTAAGATGTCGAAATATGATGATGCTGAAGGAAGG 1017

RESULT 4
AF175225
LOCUS      Homo sapiens tissue-type aorta MSTP135 mRNA linear HTC 01-AUG-2003
DEFINITION
ACCESSION AF175225.1 GI:33338055
VERSION
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM

REFERENCE
AUTHORS      Zhao, B., Xu, H.S., Tong, Y.K., Sheng, H., Qin, B.M., Liu, Y.Q., Liu, B.,
Wang, X.Y., Zhang, Q., Song, L., Gao, Y., Zhang, C.L., Ye, J., Ji, X.J.,
Liu, B.H., Lu, H., Chen, J.Z., Cai, M.Q., Zheng, W.Y., Teng, C.Y.,
Lin, Q., Fu, L.T., Lin, J., Gong, Q., Zhang, A.M., Gao, R.L. and Hui, R.T.
DIRECT SUBMISSION
TITLE      Submitted (04-AUG-1999) Molecular Medicine Center for
JOURNAL      Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167,
Beil Li Shi Lu, Beijing 100037, P.R. China
FEATURES
SOURCE      1..1292
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="aorta"
            142..1023
            /codon_start=1
            /product="MSTP135"
            /protein_id="AAQ13650.1"
            /db_xref="GI:33338056"
            /translation="MTOLAKORNFPPFWIQVLFLLTILLOKFLIKVLLDHCWGCS
            TSPHSGTGLERNILLXLFYGGPKVOLVNNRKGVKFFELNTLPLGLVYVXGIDNRGS
            CPGRLKFEAGFKRMQGIETDDQVEGLQYLASRYDFLDLRVGLHNSYGGYLSMAL
            MORSIDFVALAGAPVTLWIFPYDTGYTERYMGHPDQEGYVLGSVANQAEKFFSEPN
            RULLLHGLFDENVHFATLSILLSFLVRAGKPYDLQIYQERHSIRKVPESGHEYLHLL
            HYLOENLGSRIAALKVI"

ORIGIN
Query Match      25.1%; Score 782; DB 6; Length 1292;

```

```

Best Local Similarity 100.0%; Pred. No. 0;
Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      2325 TAAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAGAGACTTCAATATCTAGCTTCTCG 2384
Db      486 TAAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAGAGACTTCAATATCTAGCTTCTCG 545
Qy      2385 ATATGATTTTCATTGACTCGTGGTGGGATCCACGGCTGGTCTTATGGAGGATACCT 2444
Db      546 ATATGATTTTCATTGACTCGTGGGATCCACGGCTGGTCTTATGGAGGATACCT 605
Qy      2445 CTCCTCTGATGGCATTAAATGACAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCC 2504
Db      606 CTCCTCTGATGGCATTAAATGACAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCC 665
Qy      2505 AGTCACTCTGTGGATCTCTTATGATACAGGATACAGGAACTGATATATGGGTCACTGTA 2564
Db      666 AGTCACTCTGTGGATCTCTTATGATACAGGATACAGGAACTGATATATGGGTCACTGTA 725
Qy      2565 CCAGATCAACACAGGCTTATTACTTAGATCTGTGGCCATGCAAGCAGAGAAAGTTCCCTCTC 2624
Db      726 CCAGATCAACACAGGCTTATTACTTAGATCTGTGGCCATGCAAGCAGAGAAAGTTCCCTCTC 785
Qy      2625 TGAACCAATCGTTTACTGCTCTTACATGCTTTCTGGATGAGAAATGTCATTTTGCACTA 2684
Db      786 TGAACCAATCGTTTACTGCTCTTACATGCTTTCTGGATGAGAAATGTCATTTTGCACTA 845
Qy      2685 TACCAGTATATTACTGATTTTATGAGGCTGGAAAGCCATATGATTTTACAGATCTA 2744
Db      846 TACCAGTATATTACTGATTTTATGAGGCTGGAAAGCCATATGATTTTACAGATCTA 905
Qy      2745 TCCTCAGAGAGACACAGATTAAGATTCCTGAATCGGAGAAACATTAATGAATGCACTCT 2804
Db      906 TCCTCAGAGAGACACAGATTAAGATTCCTGAATCGGAGAAACATTAATGAATGCACTCT 965
Qy      2805 TTGCACTACCTTCAAGAAAACCTTTGGATCAGTATTCCTGCTCTCTAAAAGTGATATAAT 2864
Db      966 TTGCACTACCTTCAAGAAAACCTTTGGATCAGTATTCCTGCTCTCTAAAAGTGATATAAT 1025
Qy      2865 TTGACCTGTGTGAACCTCTCTGTATACATCTGGCTATTAAACCAAAATGAGAGGTTAAT 2924
Db      1026 TTGACCTGTGTGAACCTCTCTGTATACATCTGGCTATTAAACCAAAATGAGAGGTTAAT 1085
Qy      2925 CAACAGAAAACACAGAAATTTGATCATCAATTTTATACCTGCCATGTAACTCTACTCT 2984
Db      1086 CAACAGAAAACACAGAAATTTGATCATCAATTTTATACCTGCCATGTAACTCTACTCT 1145
Qy      2985 GAAAATAAAATGTGGTGCCATGCAGGGGTCTACGGTTTGTGTAGTAACTCTAACTCTTAA 3044
Db      1146 GAAAATAAAATGTGGTGCCATGCAGGGGTCTACGGTTTGTGTAGTAACTCTAACTCTTAA 1205
Qy      3045 CCCACATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACATAGAA 3104
Db      1206 CCCACATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACATAGAA 1265
Qy      3105 TT 3106
Db      1266 TT 1267

```

RESULT 5

BM557438

LOCUS

DEFINITION

AGENCY 6578992 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5466759

5', mRNA sequence.

BM557438

BM557438.1 GI:18799430

EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

(human)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

Double-stranded cDNA was ligated to Sali adaptors,
digested with NotI and cloned into the NotI and Sali sites
of the pOE90LSN_cloned vector"

ORIGIN	Query Match	Best Local Similarity	24.2%	Score 755;	DB 8;	Length 884;	
	Matches 805;	Conservative	0;	Mismatches	1;	Indels	0; Gaps 0;
Qy	25	GGCCGCTGCTACTGCGCGCTGCTTCTTAGTCGCGGTTCCGCGCTGGGTGTCACCG	84				
Db	1	GGCCGCTGCTACTGCGCGCTGCTTCTTAGTCGCGGTTCCGCGCTGGGTGTCACCG	60				
Qy	85	GGCCGCTGCTACTGCGCGCTGCTTCTTAGTCGCGGTTCCGCGCTGGGTGTCACCG	144				
Db	61	GGCCGCTGCTACTGCGCGCTGCTTCTTAGTCGCGGTTCCGCGCTGGGTGTCACCG	120				
Qy	145	GGCCGCTGCTACTGCGCGCTGCTTCTTAGTCGCGGTTCCGCGCTGGGTGTCACCG	204				
Db	121	GGCCGCTGCTACTGCGCGCTGCTTCTTAGTCGCGGTTCCGCGCTGGGTGTCACCG	180				
Qy	205	AAATGCAACATGCGAGGAGGACCTGCAACAGGACCGAGTGGAGCGCGCAGCATGAAG	264				
Db	181	AAATGCAACATGCGAGGAGGACCTGCAACAGGACCGAGTGGAGCGCGCAGCATGAAG	240				
Qy	265	GGGCACTGTGAGGAGGATATGATACAGGATCGGCTAAATGGAGCCTTTTATGTT	324				
Db	241	GGGCACTGTGAGGAGGATATGATACAGGATCGGCTAAATGGAGCCTTTTATGTT	300				
Qy	325	GAGCGGTATTCCTGGAGTCAAGTAAAGAGTGTGCGGATACCAAGAAATATCATGCG	384				
Db	301	GAGCGGTATTCCTGGAGTCAAGTAAAGAGTGTGCGGATACCAAGAAATATCATGCG	360				
Qy	385	TACATGATGCTAAGCACCACATGATTTTATGTTTGAAGAGGATGATCCAGATGGA	444				
Db	361	TACATGATGCTAAGCACCACATGATTTTATGTTTGAAGAGGATGATCCAGATGGA	420				
Qy	445	CCTCATTCAGACAGATCTATTACCTTGCATGCTGCTGAGACAGAGAAATACATG	504				
Db	421	CCTCATTCAGACAGATCTATTACCTTGCATGCTGCTGAGACAGAGAAATACATG	480				
Qy	505	TTTTATTCGAAATCCCAAACTATCAATAGACGAGCTCTTATGCTCTCTTGGAG	564				
Db	481	TTTTATTCGAAATCCCAAACTATCAATAGACGAGCTCTTATGCTCTCTTGGAG	540				
Qy	565	CCTCTTTTGGATCTTTTTCAGGCAACTGCTGATGGAATGATTTCTGAGAGAGAA	624				
Db	541	CCTCTTTTGGATCTTTTTCAGGCAACTGCTGATGGAATGATTTCTGAGAGAGAA	600				
Qy	625	CTATTAAG	684				
Db	601	CTATTAAG	660				
Qy	685	GGAAGTGGAACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTAAAGATGGAGG	744				
Db	661	GGAAGTGGAACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTAAAGATGGAGG	720				
Qy	745	CCACAGGATTTACGACAACTTTAAGCCCAATCTAGTGGAAATCTAGTTCCTCAAC	804				
Db	721	CCACAGGATTTACGACAACTTTAAGCCCAATCTAGTGGAAATCTAGTTCCTCAAC	780				
Qy	805	ATACGATGATCCCAATTTATGCCC	830				
Db	781	ATACGATGATCCCAATTTATGCCC	806				

RESULT 7
BG479035
LOCUS
DEFINITION
mRNA sequence.
EST 21-MAR-2001
BG479035
ACCESSION
BG479035.1 GI:13411314

EST.
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

EST.
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

Query Match
Best Local Similarity 100.0%; Pred. No. 0;
Matches 726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2193 GGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCGCTTGAATACCTTAGCTCTCTAGG 2252
Db 1 GGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCGCTTGAATACCTTAGCTCTCTAGG 60
Qy 2253 TTATGCTGTTGTAGTGTATAGACAAACAGGGGATCTGTCAACGAGGCTTAAATTTGAAGG 2312
Db 61 TTATGCTGTTGTAGTGTATAGACAAACAGGGGATCTGTCAACGAGGCTTAAATTTGAAGG 120
Qy 2313 CGCCTTTAAATATAAATCGGTCAAATAGAAATGACGATCAGGTGGAAGGACTCCCAATA 2372
Db 121 CGCCTTTAAATATAAATCGGTCAAATAGAAATGACGATCAGGTGGAAGGACTCCCAATA 180
Qy 2373 TCTAGCTTCTCGATATGATTTTCAATTTGACTTAGATCGTGTGGGATCCACGCTGGTCTTA 2432
Db 181 TCTAGCTTCTCGATATGATTTTCAATTTGACTTAGATCGTGTGGGATCCACGCTGGTCTTA 240
Qy 2433 TGGAGGATACCTCTCCCTGATGGCATTAAATGAGGATTAATGAGGATCAGATCTTCAAGGTTGCTAT 2492
Db 241 TGGAGGATACCTCTCCCTGATGGCATTAAATGAGGATTAATGAGGATCAGATCTTCAAGGTTGCTAT 300
Qy 2493 TGTGGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATAT 2552
Db 301 TGTGGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATAT 360
Qy 2553 GGGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATCCACGACGAGA 2612
Db 361 GGGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATCCACGACGAGA 420
Qy 2613 AAAGTTCCCTCTCGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGATGAGATGT 2672

[illegible]

COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov	
	High quality sequence stop: 630. Location/Qualifiers 1..968 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6255645" /tissue_type="epidermoid carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_102" /notes="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
FEATURES	source	
	RESULT 10 BO675260 LOCUS DEFINITION BO675260 910 bp mRNA linear EST 15-JUL-2002 AGENCOURT_8354972 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6275478 5' mRNA sequence. BO675260 VERSION BO675260.1 GI:21786094 KEYWORDS EST. SOURCE Homo sapiens (human) ORGANISM Homo sapiens Mammalia; Euteleostomi; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 910) NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCW2458 row: 1 column: 07 High quality sequence stop: 618. Location/Qualifiers 1..910 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6275478" /tissue_type="epidermoid carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_102" /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
ORIGIN	Query Match 23.2%; Score 724; DB 3; Length 968; Best Local Similarity 99.9%; Pred. No. 0; Matches 774; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
	QY 1634 AATATAACGATCCAGTGGTGGCTGCTGCTCCAGTGATTTCAAGTGTCTTATCAAG 1693 DB 1 AATATAACGATCCAGTGGTGGCTGCTGCTCCAGTGATTTCAAGTGTCTTATCAAG 60 QY 1694 AGAGATAGCAATACCACTGCTGATCGGAAGTCTTGGCCGCGATGATCAATATCC 1753 DB 61 AGAGATAGCAATACCACTGCTGATCGGAAGTCTTGGCCGCGATGATCAATATCC 120 QY 1754 AGCTTCATGATCAGAGGCTGGTATATTTTGAAGGCACAAAGATCCCTTTAGAGC 1813 DB 121 AGCTTCATGATCAGAGGCTGGTATATTTTGAAGGCACAAAGATCCCTTTAGAGC 180 QY 1814 ATCACTGTACGTAGTCACTACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTG 1873 DB 181 ATCACTGTACGTAGTCACTACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTG 240 QY 1874 GCTACTCAATCTTGTGTCATCAGTCAGCAGCTGTGCTTTTATAAGTAAAGTATAGTA 1933 DB 241 GCTACTCAATCTTGTGTCATCAGTCAGCAGCTGTGCTTTTATAAGTAAAGTATAGTA 300 QY 1934 ACCAGAAGATCCACACTGTGTGCCCTTTTACAAGCTATCAAGTCTCGAAGATGCCAA 1993 DB 301 ACCAGAAGATCCACACTGTGTGCCCTTTTACAAGCTATCAAGTCTCGAAGATGCCAA 360 QY 1994 CTTGCAAAACAAAGAAATTTTGGGCCACCAATTTTGAATTCAGCAGGTCTCTTCTGACT 2053 DB 361 CTTGCAAAACAAAGAAATTTTGGGCCACCAATTTTGAATTCAGCAGGTCTCTTCTGACT 420 QY 2054 ATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTATGATGATGC 2113 DB 421 ATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTATGATGATGC 480 QY 2114 TCTACAAGCCTCATGATCTACAGCTCGAAGAAATATCTCTGCTGTTCATATATG 2173 DB 481 TCTACAAGCCTCATGATCTACAGCTCGAAGAAATATCTCTGCTGTTCATATATG 540 QY 2174 GTGGTCTCAGGTGAGTGGTGAATATCGTTTAAAGAGTCAAGTATTCGCTTGA 2233 DB 541 GTGGTCTCAGGTGAGTGGTGAATATCGTTTAAAGAGTCAAGTATTCGCTTGA 600	
QY	2234 ATACCTAGCCTCTCTAGGTTATGTTGTTAGTAGATAGACAAACAGGGGATCCTGTACC 2293 DB 601 ATACCTAGCCTCTCTAGGTTATGTTGTTAGTAGATAGACAAACAGGGGATCCTGTACC 660 QY 2294 GAGGGCTTAAATTTGAAGCGGCTTTAAATATAAATGAGTCAATAGAAATGACGATC 2353 DB 661 GAGGGCTTAAATTTGAAGCGGCTTTAAATATAAATGAGTCAATAGAAATGACGATC 720 QY 2354 AGGTGAAGGACTCCCAATATCTAGCTTCGATATGATTTCAATGACTTAGATCG 2408 DB 721 AGGTGAAGGACTCCCAATATCTAGCTTCGATATGATTTCAATGACTTAGATCG 775	
QY	1635 ATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTTATCAAGA 1694 DB 1 ATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTTATCAAGA 60 QY 1695 GGAGATAGCAATTAACAGTGGTGAATGGGAAGTCTTGGCCGCGATGATCTAATATCCA 1754 DB 61 GGAGATAGCAATTAACAGTGGTGAATGGGAAGTCTTGGCCGCGATGATCTAATATCCA 120	

Qy	1755	AGTTGATGAGTCTGATATATTTTGAAGGACCAAGACCTCCCTTTAGAGCA	1814	/organism="Homo sapiens"
Db	121	AGTTGATGAGTCTGATATATTTTGAAGGACCAAGACCTCCCTTTAGAGCA	180	/mol_type="mRNA"
Qy	1815	TCACCTGTACGTAGTACGTAAATCTCTGGAGAGGTGACAAGGCTGACGCGTGG	1874	/db_xref="taxon:9606"
Db	181	TCACCTGTACGTAGTACGTAAATCTCTGGAGAGGTGACAAGGCTGACGCGTGG	240	/tissue_type="aorta"
Qy	1875	CTACTCACATTTCTGCTGATCAGTCAGCAGTCTCTTTTATAGTAAGTATAGTAA	1934	268..741
Db	241	CTACTCACATTTCTGCTGATCAGTCAGCAGTCTCTTTTATAGTAAGTATAGTAA	300	/codon_start=1
Qy	1935	CCAGAAGAAATCCACACTGTGTCTCTTTTACAAGCTATCAAGTCTCGAAGATGACCAAC	1994	/product="MSTP141"
Db	301	CCAGAAGAAATCCACACTGTGTCTCTTTTACAAGCTATCAAGTCTCGAAGATGACCAAC	360	/protein_id="AAQ13657.1"
Qy	1995	TTGCAAAACAAAGGAATTTTGGGCCACCAATTTTGAATTCAGCAGTCTCTTCTGACTA	2054	/db_xref="GI:33338070"
Db	361	TTGCAAAACAAAGGAATTTTGGGCCACCAATTTTGAATTCAGCAGTCTCTTCTGACTA	420	/translation="MLYKPHDLDPGKYPTVLFIYGGPOVOLVNNRFGVKYPRNLTL
Qy	2055	TACTCTCTCCAGAAATTTTCTCTTTTGAAAGTACTACTGGAATTTACATTTGATGGGATGCT	2114	ASGYVVVIDNRGSGCHLKFEGAFKYMGOIEIDQVEGLQYLASRYDFIDLDRVG
Db	421	TACTCTCTCCAGAAATTTTCTCTTTTGAAAGTACTACTGGAATTTACATTTGATGGGATGCT	480	THGWSYGYLSLMALMQRSDFRVAIAGAPVTLWIFDYDTGTYERYMGHPDQEGG"
Qy	2115	CTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTGCTGCTCATATATGG	2174	
Db	481	CTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTGCTGCTCATATATGG	540	
Qy	2175	TGGTCCCTCAGGTGCGAGTTGGTGAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAA	2234	
Db	541	TGGTCCCTCAGGTGCGAGTTGGTGAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAA	600	
Qy	2235	TACCCTAGCCTCTCTAGGTTATGCTGTGATGATAGACAAAGGGGATCCTGTACCG	2294	
Db	601	TACCCTAGCCTCTCTAGGTTATGCTGTGATGATAGACAAAGGGGATCCTGTACCG	660	
Qy	2295	AGGGCTTAAATTTGAAGGGCCCTTTAAATATAAATGGGTCAAAATAGAAATGACGATCA	2354	
Db	661	AGGGCTTAAATTTGAAGGGCCCTTTAAATATAAATGGGTCAAAATAGAAATGACGATCA	720	
Qy	2355	GGTGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTTGATAGTGTGGG	2414	
Db	721	GGTGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTTGATAGTGTGGG	780	
Qy	2415	CATCCACGGCTGGTCTCTAGGAGTACCTCTCCCTGATGGCATT	2459	
Db	781	CATCCACGGCTGGTCTCTAGGAGTACCTCTCCCTGATGGCATT	825	
RESULT 11				
AF176779				
LOCUS	AF176779	1265 bp	mRNA	linear
DEFINITION	Homo sapiens MSTP141 mRNA, complete cds.			
ACCESSION	AF176779			
VERSION	AF176779.1	GI:33338069		
KEYWORDS	HTC.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;				
Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1265)			
AUTHORS	Hui,R.T., Liu,Y.Q., Wang,X.Y., Qin,B.M. and Sheng,H.			
TITLE	Homo sapiens normal aorta mRNA MST141			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 1265)			
AUTHORS	Hui,R.T., Liu,Y.Q., Wang,X.Y., Qin,B.M. and Sheng,H.			
TITLE	Direct Submission			
JOURNAL	Submitted (10-AUG-1999) Molecular Medicine Center for			
Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167,				
Bei Li Shi Lu, Beijing 100037, P.R. China				
FEATURES	Location/Qualifiers			
1..1265				
source				

Qy	1861	CTGACTGACCGTGGCTACTCACATCTCTGCTGATCAGTCAGCAGTCTTCTTTATA	1920
Db	19	CTGACTGACCGTGGCTACTCACATCTCTGCTGATCAGTCAGCAGTCTTCTTTATA	78
Qy	1921	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCTCTTTACAAGCTATCAAGTCT	1980
Db	79	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCTCTTTACAAGCTATCAAGTCT	138
Qy	1981	GAAGATGACCACTTGCRAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGT	2040
Db	139	GAAGATGACCACTTGCRAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGT	198
Qy	2041	CCTCTTCTGACTATATCTCTCCAGAAAATTTTCTCTTTTGAAGTACTACTCGATTTACA	2100
Db	199	CCTCTTCTGACTATATCTCTCCAGAAAATTTTCTCTTTTGAAGTACTACTCGATTTACA	258
Qy	2101	TTGTATCGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTG	2160
Db	259	TTGTATCGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTG	318
Qy	2161	CTGTTCAATATATGGTGGTCTCTCAGGTGCGAGTTGGTGAATTAATCGGTTTAAAGAGTCAAG	2220
Db	319	CTGTTCAATATATGGTGGTCTCTCAGGTGCGAGTTGGTGAATTAATCGGTTTAAAGAGTCAAG	378
Qy	2221	TATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGCTGTGTAGTATAGACAACAGG	2280
Db	379	TATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGCTGTGTAGTATAGACAACAGG	438
Qy	2281	GGATCTCTGTCACCGAGGCTTAAATTTGAAGGGCCCTTTAAATATAAATGGGTCAATA	2340
Db	439	GGATCTCTGTCACCGAGGCTTAAATTTGAAGGGCCCTTTAAATATAAATGGGTCAATA	498
Qy	2341	GAATTCGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTCAC	2400
Db	499	GAATTCGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTCAC	558
Qy	2401	TTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTA	2460
Db	559	TTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTA	618
Qy	2461	ATGACAGAGGTCAGATATCTTTTCAAGGTTGCTATTTGCTGGGGCCCCCAGTCACTCTGTGGATC	2520
Db	619	ATGACAGAGGTCAGATATCTTTTCAAGGTTGCTATTTGCTGGGGCCCCCAGTCACTCTGTGGATC	678
Qy	2521	TTCTATGATACAGGATACAGGAACTTATATGGGTCACTCCCTGACCAAGATGACAGGCG	2580
Db	679	TTCTATGATACAGGATACAGGAACTTATATGGGTCACTCCCTGACCAAGATGACAGGCG	738
Qy	2581	TA 2582	
Db	739	TA 740	
RESULT 12			
AL043338			

CDS				/organism="Homo sapiens"			
				/mol_type="mRNA"			
				/db_xref="taxon:9606"			
				/tissue_type="aorta"			
				268..741			
				/codon_start=1			
				/product="MSTP141"			
				/protein_id="AAQ13657.1"			
				/db_xref="GI:33338070"			
				/translation="MLYKPHDLDPGKYPTVLFIYGGPOVOLVNNRFGVKYPRNLTL			
				ASGYVVVIDNRGSGCHLKFEGAFKYMGOIEIDQVEGLQYLASRYDFIDLDRVG			
				THGWSYGYLSLMALMQRSDFRVAIAGAPVTLWIFDYDTGTYERYMGHPDQEGG"			
ORIGIN							
				Query Match	23.1%;	Score 722;	DB 6; Length 1265;
				Best Local Similarity	100.0%;	Pred. No. 0;	
				Mismatches	722;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1861	CTGACTGACCGTGGCTACTCACAATTTCTGCTGCATCAGTCAGCAGTCTCTTTATA	1920				
Db	19	CTGACTGACCGTGGCTACTCACAATTTCTGCTGCATCAGTCAGCAGTCTCTTTATA	78				
Qy	1921	AGTAAGTATATAGTAACCAAGAAATCCACACTGTGTGCTTTCACAGCTATCAAGTCT	1980				
Db	79	AGTAAGTATATAGTAACCAAGAAATCCACACTGTGTGCTTTCACAGCTATCAAGTCT	138				
Qy	1981	GAAGATGACCAACTTGCAAAACAAAGAAATTTGGGCCACCACTTTGGATTCAGCAGT	2040				
Db	139	GAAGATGACCAACTTGCAAAACAAAGAAATTTGGGCCACCACTTTGGATTCAGCAGT	198				
Qy	2041	CCTCTCTCAGCTATATCTCTCCAGAAAATTTCTCTTTTGAAGTACTACTGGATTTACA	2100				
Db	199	CCTCTCTCAGCTATATCTCTCCAGAAAATTTCTCTTTTGAAGTACTACTGGATTTACA	258				
Qy	2101	TTGTATGGGATGCTCTACAGCCTCANGATCTACAGCCTGGAAAGAAATATCTACTGTG	2160				
Db	259	TTGTATGGGATGCTCTACAGCCTCANGATCTACAGCCTGGAAAGAAATATCTACTGTG	318				
Qy	2161	CTGTTTCAATATATGCTGCTCTCAGGTGAGTGGTGAATAATCGGTTTAAAGGAGTCAAG	2220				
Db	319	CTGTTTCAATATATGCTGCTCTCAGGTGAGTGGTGAATAATCGGTTTAAAGGAGTCAAG	378				
Qy	2221	TATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTATGCTGTTGTAGTATAGACAGG	2280				
Db	379	TATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTATGCTGTTGTAGTATAGACAGG	438				
Qy	2281	GGATCCTGTCACCGGGCTTAAATTTGAAGGGCCCTTTAAATATAAAATGGTCAATA	2340				
Db	439	GGATCCTGTCACCGGGCTTAAATTTGAAGGGCCCTTTAAATATAAAATGGTCAATA	498				
Qy	2341	GAATTTGACGATCAGGTGGAGGATCCCAATATCTAGCTTCTCGATATGATTTCAATTGAC	2400				
Db	499	GAATTTGACGATCAGGTGGAGGATCCCAATATCTAGCTTCTCGATATGATTTCAATTGAC	558				
Qy	2401	TTAGATCGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGACATTA	2460				
Db	559	TTAGATCGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGACATTA	618				
Qy	2461	ATGACAGGTCAGATATCTTTTCCAGGTTGCTTATTTGCTGGGGCCCAAGTCACTCTGTGGATC	2520				
Db	619	ATGACAGGTCAGATATCTTTTCCAGGTTGCTTATTTGCTGGGGCCCAAGTCACTCTGTGGATC	678				
Qy	2521	TTCTATGATACAGGATACACGGAACTTATATGGGTGACCTTGACCAAGTAACACAGGCG	2580				
Db	679	TTCTATGATACAGGATACACGGAACTTATATGGGTGACCTTGACCAAGTAACACAGGCG	738				
Qy	2581	TA 2582					
Db	739	TA 740					
RESULT 12							
AL043338							

LOCUS
 DEFINITION DKFp43400723_r1 434 (synonym: htes3) mRNA linear EST 04-SEP-2003
 ACCESSION DKFp43400723 5', mRNA sequence.
 VERSION AL043338
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 735)
 Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
 EST (Blum, et al.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: MIPS
 MIPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by LMU (Ludwig Maximilians University,
 Munich/Germany) within the CDNA sequencing consortium of the German
 Genome Project.
 No sl sequence available.
 This clone (DKFp43400723) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
 FEATURES
 source
 1..735
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFp43400723"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="434 (synonym: htes3)"
 /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN
 Query Match 22.7%; Score 708; DB 1; Length 735;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2274 CAACAGGGGATCTGTGTCACCGAGGGCTTAAATTGAGGCGCTTTAAATATTAATATAAATGGG 2333
 DB 1 CAACAGGGGATCTGTGTCACCGAGGGCTTAAATTGAGGCGCTTTAAATATTAATATAAATGGG 60
 QY 2334 TCAATAGAAATTGACGATCAGGTGAAGGACTCCAATATCTAGCTTCGTGATATGATTT 2393
 DB 61 TCAATAGAAATTGACGATCAGGTGAAGGACTCCAATATCTAGCTTCGTGATATGATTT 120
 QY 2394 CATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCCTATGAGGATACCTCTCCCTGAT 2453
 DB 121 CATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCCTATGAGGATACCTCTCCCTGAT 180
 QY 2454 GGCAATTAATGACAGGTTCAGATATCTTCAGGGTGTCTATTCCTGGGGCCCACTCTCT 2513
 DB 181 GGCAATTAATGACAGGTTCAGATATCTTCAGGGTGTCTATTCCTGGGGCCCACTCTCT 240
 QY 2514 GTGGATCTTCTATGATACAGGATACACCGAGGTATATGGTCACTCGACAGAAATCA 2573
 DB 241 GTGGATCTTCTATGATACAGGATACACCGAGGTATATGGTCACTCGACAGAAATCA 300
 QY 2574 ACAGGGCTATTACTTAGGATCTGGCCATCCAGCAAGAAAGTTCCTCTGAAACCAA 2633
 DB 301 ACAGGGCTATTACTTAGGATCTGGCCATCCAGCAAGAAAGTTCCTCTGAAACCAA 360
 QY 2634 TCGTTTACTGCTCTTACATGGTTCTCGATGAGATGTCCATTTTGCACATACCATGAT 2693
 DB 361 TCGTTTACTGCTCTTACATGGTTCTCGATGAGATGTCCATTTTGCACATACCATGAT 420

QY 2694 ATTACTCAGTTTTTTAGTGAGGCTCGGAAGCCATATGATTTACAGATCTATCTCTCAGGA 2753
 DB 421 ATTACTCAGTTTTTTAGTGAGGCTCGGAAGCCATATGATTTACAGATCTATCTCTCAGGA 480
 QY 2754 GAGACACAGCATAAGAGATTCCTGAAATCGGAGCAATTTATGAATCTGCACTA 2813
 DB 481 GAGACACAGCATAAGAGATTCCTGAAATCGGAGCAATTTATGAATCTGCACTA 540
 QY 2814 CTTTCAAGAAACCTTGAATCAGTATTTGCTGCTCTTAAAGTATATAATTTGACCTGT 2873
 DB 541 CTTTCAAGAAACCTTGAATCAGTATTTGCTGCTCTTAAAGTATATAATTTGACCTGT 600
 QY 2874 GTAGAACTCTCGGTATACACGCTGCTATTTAAACAAATGAGGAGTTTAAATCAACAGAAA 2933
 DB 601 GTAGAACTCTCGGTATACACGCTGCTATTTAAACAAATGAGGAGTTTAAATCAACAGAAA 660
 QY 2934 ACACAGAAATGATCATCATTTTGAATCCTGCCATGTAACATCTACT 2981
 DB 661 ACACAGAAATGATCATCATTTTGAATCCTGCCATGTAACATCTACT 708
 RESULT 13
 LOCUS AL040398
 DEFINITION DKFp434A0714_r1 434 (synonym: htes3) mRNA linear EST 04-SEP-2003
 ACCESSION DKFp434A0714 5', mRNA sequence.
 VERSION AL040398
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 753)
 Koehrer, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
 EST (Koehrer, et al.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: MIPS
 MIPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by BMFZ (Biomedical Research Center at the Charite,
 Berlin/Germany) within the CDNA sequencing consortium of the German
 Genome Project.
 No sl sequence available.
 This clone (DKFp434A0714) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
 FEATURES
 source
 1..753
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFp434A0714"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="434 (synonym: htes3)"
 /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
 ORIGIN
 Query Match 22.4%; Score 700; DB 1; Length 753;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 750; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1609 ATTACATCTATTTAAAGGAAAGCAATATAAAGTTCAGTGGGCTGCTCTCCA 1668
 DB 1 ATTACATCTATTTAAAGGAAAGCAATATAAAGTTCAGTGGGCTGCTCTCCA 60
 QY 1669 AGTGATTTCAAGTCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGAGT 1728


```

Db      61 AGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTTACCAAGTGGTGAATGGGAAGTT 120
Qy      1729 CTTGGCGCGCATGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGATATATTTTGA 1788
Db      121 CTTGGCGCGCATGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGATATATTTTGA 180
Qy      1789 GGCACCAAGAGCTCCCTTTAGAGCATCCCTGTAGTAGTACGTTAGTAAATCTCTGGA 1848
Db      181 GGCACCAAGAGCTCCCTTTAGAGCATCCCTGTAGTAGTACGTTAGTAAATCTCTGGA 240
Qy      1849 GAGGTGACCAAGGCTGACGACCTGCTACTACATTTCTGTGTCATCAGTCAGCACTGT 1908
Db      241 GAGGTGACCAAGGCTGACGACCTGCTACTACATTTCTGTGTCATCAGTCAGCACTGT 300
Qy      1909 GACTTCTTTAAGTAGTAGTATAGTAACCAAGAGATCCACACTGTGTCTCTTTTACAAG 1968
Db      301 GACTTCTTTAAGTAGTAGTATAGTAACCAAGAGATCCACACTGTGTCTCTTTTACAAG 360
Qy      1969 CTATCAAGTCTGAGATGACCAACTTGCACCAACAAAGGAATTTTGGGCCACCAATTTG 2028
Db      361 CTATCAAGTCTGAGATGACCAACTTGCACCAACAAAGGAATTTTGGGCCACCAATTTG 420
Qy      2029 GATTGACAGGTCCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTTGAAGTACT 2088
Db      421 GATTGACAGGTCCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTTGAAGTACT 480
Qy      2089 ACTGGATTTACATTTGATGGATGCTCTACAGGCTCATGATACAGCTTACAGCTTGAAGAAA 2148
Db      481 ACTGGATTTACATTTGATGGATGCTCTACAGGCTCATGATACAGCTTGAAGAAA 540
Qy      2149 TATCCTACTGTGCTTTCAATATATATGTTGGTCTCAGGTGCAAGTGGTGAATTAATCGTTT 2208
Db      541 TATCCTACTGTGCTTTCAATATATATGTTGGTCTCAGGTGCAAGTGGTGAATTAATCGTTT 600
Qy      2209 AAAGGATCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTTGTTAGTG 2268
Db      601 AAAGGATCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTTGTTAGTG 660
Qy      2269 ATAGACAAACAGGGGATCCTGTACCGAGGGCTTAAATTTGAAGGGCCCTTTAAATATAAA 2328
Db      661 ATAGACAAACAGGGGATCCTGTACCGAGGGCTTAAATTTGAAGGGCCCTTTAAATATAAA 720
Qy      2329 ATGGGTCAATAGAAATGACATCAGGTGG 2359
Db      721 ATGGGTCAATAGAAATGACATCAGGTGG 751

```

```

RESULT 14
CR998849
LOCUS      CR998849      835 bp      mRNA      linear      EST 28-JUN-2005
DEFINITION CR998849 RZPD no. 9016 Homo sapiens cDNA clone RZPDp9016K2027 5',
mRNA sequence.
ACCESSION CR998849
VERSION   CR998849.1 GI:68292734
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 835)
Hell.O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
and Korn,B.
Human T-Lymphocytes library
Unpublished (2005)
Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDp9016K2027.
RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
http://www.rzpd.de/cgi-bin/products/aet.cgi?libNo=9016 Contact:

```

```

Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel.: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=RZPDp9016K2027
contact RZPD (product-support@rzpd.de) for further information.
Primer name: qe3_4, Primer sequence: CGGATAACAATTCACACAG.
Location/Qualifiers
1. .835
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="RZPDp9016K2027"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9016"
/note="vector: PQE80LSN_cloned; Site_1: SalI; Site_2:
NotI; vector:
http://www.rzpd.de/info/vectors/PQE80LSN_cloned_pic.shtml
; 1st strand cDNA was prepared from mRNA obtained from
human T-Lymphocytes with a NotI - oligo(dt) primer [5',
GACTAGTCTCTAGATCGGAGCGCGCCCTCTTTTTTTTTTTT 3'] .
Double-stranded cDNA was ligated to SalI adaptors,
digested with NotI and cloned into the NotI and SalI sites
of the PQE80LSN_cloned vector"

```

```

FEATURES
source
Query Match 22.1%; Score 690; DB 8; Length 835;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 740; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      78 GTCACCGCGCGCGCGCGGAGGAGCCACTGCAACACGAGCGGAGTGGAGCGCGCGAG 137
Db      1 GTCACCGCGCGCGCGCGGAGGAGCCACTGCAACACGAGCGGAGTGGAGCGCGCGAG 60
Qy      138 CATGAAGCGCGCGCGCGCGGCTCCATAGCGCAGCTCGGACGCTCGGCGCGCGCGG 197
Db      61 CATGAAGCGCGCGCGCGCGGCTCCATAGCGCAGCTCGGACGCTCGGCGCGCGCGG 120
Qy      198 GGAAGGAATGCAACATGCGCAGCAATGGAACACAGACAGCTGGGTGTTGAGATATT 257
Db      121 GGAAGGAATGCAACATGCGCAGCAATGGAACACAGACAGCTGGGTGTTGAGATATT 180
Qy      258 TGAACCTCGGAGCTGTGAGGAGAAATTTGAATCACAAGATCGGCCTAAATTGGAGCCTTT 317
Db      181 TGAACCTCGAGACTGTGAGGAGAAATTTGAATCACAAGATCGGCCTAAATTGGAGCCTTT 240
Qy      318 TTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTTCCGATACCCAGAAAATA 377
Db      241 TTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTTCCGATACCCAGAAAATA 300
Qy      378 TCATGGCTTACATGATGGCTAAGGCACCAATCATGATTTTCATGTTTGTGAAGAGAAATGATCC 437
Db      301 TCATGGCTTACATGATGGCTAAGGCACCAATCATGATTTTCATGTTTGTGAAGAGAAATGATCC 360
Qy      438 AGATGGACCTCATTCAGACAGAAATCTTACCTTCCATGCTGCTGGAGAACAGAGAAA 497
Db      361 AGATGGACCTCATTCAGACAGAAATCTTACCTTCCATGCTGCTGGAGAACAGAGAAA 420
Qy      498 TACACTGTTTATTTCTGAAAATCCCAAACTATCAATAGACAGCAGCTTAAATGCTCTC 557
Db      421 TACACTGTTTATTTCTGAAAATCCCAAACTATCAATAGACAGCAGCTTAAATGCTCTC 480
Qy      558 TTGGAAGCCTCTTTTGGATCTTTTTCAGGCACACCTGAGACTATGGAATGATTTCTCGAGA 617
Db      481 TTGGAAGCCTCTTTTGGATCTTTTTCAGGCACACCTGAGACTATGGAATGATTTCTCGAGA 540
Qy      618 AGAAGAACTATTAGAGAAAAGAAAAGCAATTCGAAACAGTCGGAATTCCTTACGATTA 677

```


Thu Jun 22 09:04:26 2006

RESULT 17
BX390898
LOCUS
DEFINITION
BX390898 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens cDNA clone CS0DL005YD02 5-PRIME, mRNA sequence.
ACCESSION
BX390898.2 GI:46844267
VERSION
BX390898
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1. (bases 1 to 914)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 13, 2003 this sequence version replaced gi:30607432.
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and Ecor V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7542.i
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAG053ZD07_CS05044_1&c=7542.i

FEATURES
source
Location/Qualifiers
1..914
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="GRN PRENEU"
/note="oligo dT primed, full-length enriched cDNA library
from hES cell line H7 (p29) maintained in feeder-free
conditions. Embryoid bodies were generated in the presence
of all-trans retinoic acid and mitogens."

ORIGIN
Query Match 22.0%; Score 687; DB 8; Length 742;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 737; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 552 GCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGATTC 611
DB 5 GCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGATTC 64
QY 612 TCGAAGAAGAACTATTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 671
DB 65 TCGAAGAAGAACTATTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 124
QY 672 CGATTATACCAAG 731
DB 125 CGATTATACCAAG 184
QY 732 AAAAGATGAGAGGCGCACAGAGATTTACGACAACTTTAAGCCCAATCTAGTGAAGC 791
DB 185 AAAAGATGAGAGGCGCACAGAGATTTACGACAACTTTAAGCCCAATCTAGTGAAGC 244
QY 792 TAGTTGTCACACATACGAGTGGATCCAAATATATGCGCGCTGATCCAGATGGAATGC 851
DB 245 TAGTTGTCACACATACGAGTGGATCCAAATATATGCGCGCTGATCCAGATGGAATGC 304
QY 852 TTTTATACATAGCAACGATATTGGATATCTAAATCGTAAACAGAGAGAGAGAGAGACT 911
DB 305 TTTTATACATAGCAACGATATTGGATATCTAAATCGTAAACAGAGAGAGAGAGACT 364
QY 912 CACTTATGTGCAATGAGCTAGCCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGC 971
DB 365 CACTTATGTGCAATGAGCTAGCCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGC 424
QY 972 TACCTTTGTTCTCCAGAGAGATTTGATAGATATTCTGGCTATTGGTGTGTCGAAAGC 1031
DB 425 TACCTTTGTTCTCCAGAGAGATTTGATAGATATTCTGGCTATTGGTGTGTCGAAAGC 484
QY 1032 TGAACAACCTCCAGTGGTGGTAAATTTTGAATTTCTATAGAAATGATGATTC 1091
DB 485 TGAACAACCTCCAGTGGTGGTAAATTTTGAATTTCTATAGAAATGATGATTC 544
QY 1092 TGAGGTGGAATTTATTCATCTACATCCCTATGTTGGAACAAGAGAGGAGAGATTCATT 1151
DB 545 TGAGGTGGAATTTATTCATCTACATCCCTATGTTGGAACAAGAGAGGAGAGATTCATT 604
QY 1152 CCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAT 1211
DB 605 CCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAT 664
QY 1212 GATTGATGCTGAAGAGGATCATAGATGTCATAGATAGAGAACTAAATCAACCTTTTGA 1271
DB 665 GATTGATGCTGAAGAGGATCATAGATGTCATAGATAGAGAACTAAATCAACCTTTTGA 724
QY 1272 GATTCTATTGAGGAGT 1289
DB 725 GATTCTATTGAGGAGT 742

QY 739 GGAGGGCCACAGGATTTACGCAACACCTTTAAGGCCCAATCTAGTGGAACTAGTTGT 798
DB 73 GGAGGGCCACAGGATTTACGCAACACCTTTAAGGCCCAATCTAGTGGAACTAGTTGT 132
QY 799 CCACAACATACGAGTGGATCCAAATATGCGCGCTGATCCAGACTGGATTCGTTTATA 858
DB 133 CCACAACATACGAGTGGATCCAAATATGCGCGCTGATCCAGACTGGATTCGTTTATA 192
QY 859 CATAGCAACGATTTTGGATATCTAAATCGTAAACAGAGAGAGAGAGACTCATTAT 918
DB 193 CATAGCAACGATTTTGGATATCTAAATCGTAAACAGAGAGAGAGAGACTCATTAT 252
QY 919 GTGCACAATAGCTAGCTAGCCCAACATGGAAGAGATGCCAGATCAGCTGAGTACCTTT 978
DB 253 GTGCACAATAGCTAGCTAGCCCAACATGGAAGAGATGCCAGATCAGCTGAGTACCTTT 312
QY 979 GTTCTCCAGAGAGATTTGATAGATATTCTGGCTATTGGTGTGTCGAAAGCTGAACA 1038
DB 313 GTTCTCCAGAGAGATTTGATAGATATTCTGGCTATTGGTGTGTCGAAAGCTGAACA 372
QY 1039 ACTCCAGTGGTGGTAAATTTCTAGATTTCTATAGAGAGAGAGAGAGAGAGAGAG 1098

373 ACTCCAGTGGTGTAAATTTCTTAGAATCTTATGAAGAAAATGATGAATCTGAGTG 432
 1099 GAATATTTATGTTATCCCTATGTTGGAACAAGGAGGCGAGATTCATTCGGTTAT 1158
 433 GAATATTTATGTTATCCCTATGTTGGAACAAGGAGGCGAGATTCATTCGGTTAT 492
 1159 CCTAAACAGGTACAGCAATCCCTAAAGTCACTTTAAAGATGTCAGAAATATGATTCAT 1218
 493 CCTAAACAGGTACAGCAATCCCTAAAGTCACTTTAAAGATGTCAGAAATATGATTCAT 552
 1219 GCTGAAGGAAGATCATAGATGTCATAGATAAGCAACTAATTCACCTTTTGAATCTTA 1278
 553 GCTGAAGGAAGATCATAGATGTCATAGATAAGCAACTAATTCACCTTTTGAATCTTA 612
 1279 TTTGAAGGAGTGAATATATTTGCCAGAGCTGAGTGAATCTTGGGGAATATGCTTGG 1338
 613 TTTGAAGGAGTGAATATATTTGCCAGAGCTGAGTGAATCTTGGGGAATATGCTTGG 672
 1339 TCCATCTCTACTAGATGCTCCAGACTCGCTTACAGATAGTGTGATCTACCTGAATTA 1398
 673 TCCATCTCTACTAGATGCTCCAGACTCGCTTACAGATAGTGTGATCTACCTGAATTA 732
 1399 TTTATCCAGTGAAGATGATGTTATGGAAGGAGGAGCACTATGATGTCAGTGCCTGAT 1458
 733 TTTATCCAGTGAAGATGATGTTATGGAAGGAGGAGCACTATGATGTCAGTGCCTGAT 792
 1459 TCTGTGACGCCACTAAT 1475
 793 TCTGTGACGCCACTAAT 809

RESULT 18
 CK000768
 LOCUS
 DEFINITION
 AGENCOURT 16368692 NIH MGC 221 Homo sapiens cDNA clone
 IMAGE:30707852 5', mRNA sequence.
 CK000768
 CK000768.1 GI:38526802
 EST.
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo; 724
 1 (bases) 1 (bp) 724
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute
 Bldg 31 Rm10A07 Bethesda, MD 20892
 Email: cga@os-xemail.nih.gov
 Tissue Procurement: James Martin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDAMI073 row: b column: 21
 High quality sequence stop: 678.
 Location/Qualifiers
 1..724
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30707852"
 /lab_host="DH10B ToaA"
 /clone_lib="NIH MGC 221"
 /note="Organ: mixed; Vector: pX-Asc; Site 1: EcoRI;
 Site 2: NotI; Library is oligo-dT primed and directionally
 cloned. Denatured RNA was size fractionated on a 1% agarose

FEATURES

source

RESULT 19
 CK994557
 LOCUS
 DEFINITION
 ACCESSION

CR994557 820 bp mRNA linear EST 28-JUN-2005
 CR994557 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016A0619 5',
 mRNA sequence.
 CR994557

gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to RNA size fraction, ligated with
 EcoR I adaptor digested with Not I and then cloned
 directionally into pX-Asc vector. Average insert size
 4-5kb. Adaptors 5', 3', linker sequence - GCGGCGCTGAGACC T18.
 Sequencing primers 3' end: T3 promoter primer 5'd
 (ATTACCTCTACTAAGGAG)3'. 5' End: T7 promoter primer 5'd
 (TAATAGGCTACTATAGG)3'. Library was constructed in the
 laboratory of M. Bento Soares. Note: this is a NIH_MGC
 Library"

ORIGIN

Query Match 22.0%; Score 685; DB 5; Length 724;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 685; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 23 AAGCGCGCTCTACTCCGCGCTGCTTTAGTCCGCGTTCGCCCGCTGGGTGTGCAC 82
 Db 8 AAGCGCGCTCTACTCCGCGCTGCTTTAGTCCGCGTTCGCCCGCTGGGTGTGCAC 67
 QY 83 CGCGCGCGCGCGCGAGGAGGAGCACTGCAACAGGACCGGAGTGGAGCGCGCGAGATGA 142
 Db 68 CGCGCGCGCGCGCGAGGAGGAGCACTGCAACAGGACCGGAGTGGAGCGCGCGAGATGA 127
 QY 143 AAGCGCGCGCGCGCTCCATAGCGCGCTCCGCGCGCGTCCGCGCGCGCGCGCGGAG 202
 Db 128 AAGCGCGCGCGCGCTCCATAGCGCGCTCCGCGCGCGTCCGCGCGCGCGCGCGGAG 187
 QY 203 GAAATGCAACATGCGAGCAATGAAACAGAACAGCTGGGTGTGAGATATTGAAA 262
 Db 188 GAAATGCAACATGCGAGCAATGAAACAGAACAGCTGGGTGTGAGATATTGAAA 247
 QY 263 CTGCGGACTGTGAGGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 322
 Db 248 CTGCGGACTGTGAGGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 307
 QY 323 TTGAGCGGTATTCCTGGAGTCACTTAAAGAGTCTTCCGCGCGTCCGCGCGTCCGCGCG 382
 Db 308 TTGAGCGGTATTCCTGGAGTCACTTAAAGAGTCTTCCGCGCGTCCGCGCGTCCGCGCG 367
 QY 383 GCTACATGATGGCTTAAGGACCAACATGATTTTCAATTTTGTGAGAGGATGATCCAGATG 442
 Db 368 GCTACATGATGGCTTAAGGACCAACATGATTTTCAATTTTGTGAGAGGATGATCCAGATG 427
 QY 443 GACCTATTGAGACAGATCTTATACCTTGGCCATGCTGCTGAGGAGGAGGAGGAGGAG 502
 Db 428 GACCTATTGAGACAGATCTTATACCTTGGCCATGCTGCTGAGGAGGAGGAGGAGGAG 487
 QY 503 TGTATTATCTGAAATTTCCAAACTATCAATAGAGGAGGAGGAGGAGGAGGAGGAGGAG 562
 Db 488 TGTATTATCTGAAATTTCCAAACTATCAATAGAGGAGGAGGAGGAGGAGGAGGAGGAG 547
 QY 563 AGCTCTCTTTGGATCTTTTTCAGGCAACATGATGATGATGATGATGATGATGATGATGAT 622
 Db 548 AGCTCTCTTTGGATCTTTTTCAGGCAACATGATGATGATGATGATGATGATGATGATGAT 607
 QY 623 AACTATTAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 682
 Db 608 AACTATTAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 667
 QY 683 AAGGAGTGGACATTTCTGTTTCA 707
 Db 668 AAGGAGTGGACATTTCTGTTTCA 692

ORIGIN

```
Query Match      21.7%; Score 676; DB 1; Length 864;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2268 GATAGACAAACAGGGGATCTGTCAACGAGGGCTTAAATTTGAAGGCGCCCTTTAAATATAA 2327
    |||||
DB 1 GATAGACAAACAGGGGATCTGTCAACGAGGGCTTAAATTTGAAGGCGCCCTTTAAATATAA 60
QY 2328 AATGGGTCAAAATAGAAATGACGATCAGGTGCGAGGACCTCAATATCTAGCTTCTCGATA 2387
    |||||
DB 61 AATGGGTCAAAATAGAAATGACGATCAGGTGCGAGGACCTCAATATCTAGCTTCTCGATA 120
QY 2388 TGATTTTCATTGACTTAGATCTGTGCGCATCCAGGCTGCTCTATGGAGGATACCTCTC 2447
    |||||
DB 121 TGATTTTCATTGACTTAGATCTGTGCGCATCCAGGCTGCTCTATGGAGGATACCTCTC 180
QY 2448 CCTGATGGCAATTAATGACAGGTCAGATATCTTCAGGTTGCTATGCTGGGGCCCCAGT 2507
    |||||
DB 181 CCTGATGGCAATTAATGACAGGTCAGATATCTTCAGGTTGCTATGCTGGGGCCCCAGT 240
QY 2508 CACTCTGTGGATCTTCTATGATACAGGATACAGGACGTTATATGGGTCAACCTGACCA 2567
    |||||
DB 241 CACTCTGTGGATCTTCTATGATACAGGATACAGGACGTTATATGGGTCAACCTGACCA 300
QY 2568 GAATGAAACAGGGCTTATCTAGTATCTGAGGCTGGAAGCCATATGATTTACAGATCTATCC 2627
    |||||
DB 301 GAATGAAACAGGGCTTATCTAGTATCTGAGGCTGGAAGCCATATGATTTACAGATCTATCC 360
QY 2628 ACCAAATCGTTTACTGCTCTTACATGCTTCTGATGAGATGAGATGTCATTTTGACATAC 2687
    |||||
DB 361 ACCAAATCGTTTACTGCTCTTACATGCTTCTGATGAGATGAGATGTCATTTTGACATAC 420
QY 2688 CAGTATATCTAGTATCTTACTGAGGCTGGAAGCCATATGATTTACAGATCTATCC 2747
    |||||
DB 421 CAGTATATCTAGTATCTTACTGAGGCTGGAAGCCATATGATTTACAGATCTATCC 480
QY 2748 TCAGGAGACACAGCATATGAGTCTCGAATCGGAGAACTATGAACTGCACTCTTTT 2807
    |||||
DB 481 TCAGGAGACACAGCATATGAGTCTCGAATCGGAGAACTATGAACTGCACTCTTTT 540
QY 2808 GCATCTACCTTCAAGAAACCTTGGATCAGTATGCTCTCTAAAGTATGATATATTTTG 2867
    |||||
DB 541 GCATCTACCTTCAAGAAACCTTGGATCAGTATGCTCTCTAAAGTATGATATATTTTG 600
QY 2868 ACTGTGTAGACTCTCTGATATACATGCTATTTAACCAATGAGGAGGTTTAAATCAA 2927
    |||||
DB 601 ACTGTGTAGACTCTCTGATATACATGCTATTTAACCAATGAGGAGGTTTAAATCAA 660
QY 2928 CAGAAACACAGAAAT 2943
    |||||
DB 661 CAGAAACACAGAAAT 676
    |||||
```

```
RESULT 21
AY411616
LOCUS
DEFINITION
Pan troglodytes DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,
AY411616 survey sequence.
ACCESSION
AY411616.1 GI:39767584
VERSION
GSS
KEYWORDS
Pan troglodytes (chimpanzee)
SOURCE
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pan.
1 (bases 1 to 2649)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
```

TITLE

JOURNAL
PUBMED
REFERENCE
AUTHORS
2 (bases 1 to 2649)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE

JOURNAL
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES

source
1..2649
location/Qualifiers
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..2649
/gene="DPP8"
/locus_tag="HCM4283"

ORIGIN

```
Query Match      21.6%; Score 674; DB 14; Length 2649;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 824; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1750 ATCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTA 1809
    |||||
DB 1537 ATCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTA 1596
QY 1810 GAGCATCACCTCTAGCTAGTCAGTTACGTAATCTCTGAGAGGTGACAGGCTGACTGAC 1869
    |||||
DB 1597 GAGCATCACCTCTAGCTAGTCAGTTACGTAATCTCTGAGAGGTGACAGGCTGACTGAC 1656
QY 1870 CBTGGCTACTCACATTTCTTCTGTCATCAGTCAGCACTGTGACTCTTTTAAAGTAAGTAT 1929
    |||||
DB 1657 CBTGGCTACTCACATTTCTTCTGTCATCAGTCAGCACTGTGACTCTTTTAAAGTAAGTAT 1716
QY 1930 AGTAAACCAAGAAATCCACACTGTGTCCTTTTACAAGCTATCAAGTCTCGAAGATGAC 1989
    |||||
DB 1717 AGTAAACCAAGAAATCCACACTGTGTCCTTTTACAAGCTATCAAGTCTCGAAGATGAC 1776
QY 1990 CCAACTTGCAAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTTCAGAGGTCCTCTTCT 2049
    |||||
DB 1777 CCAACTTGCAAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTTCAGAGGTCCTCTTCT 1836
QY 2050 GACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGG 2109
    |||||
DB 1837 GACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGG 1896
QY 2110 ATGCTCTACAAGCTCTCATGATCTACAGCTCGAAGAAATATCTCTACTGTGCTGTTTCA 2169
    |||||
DB 1897 ATGCTCTACAAGCTCTCATGATCTACAGCTCGAAGAAATATCTCTACTGTGCTGTTTCA 1956
QY 2170 TATGGTGTCTCTCAGGTGAGTTGGTGAATATCGTTTAAAGAGTCAAGTATTTCCCG 2229
    |||||
DB 1957 TATGGTGTCTCTCAGGTGAGTTGGTGAATATCGTTTAAAGAGTCAAGTATTTCCCG 2016
QY 2230 TTGAATACCTCTAGCTCTCTAGTTTCTGGTTGATGATAGACAAACAGGGATCTCTGT 2289
    |||||
DB 2017 TTGAATACCTCTAGCTCTCTAGTTTCTGGTTGATGATAGACAAACAGGGATCTCTGT 2076
QY 2290 CACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAATATAGAAATGAC 2349
    |||||
DB 2077 CACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAATATAGAAATGAC 2136
QY 2350 GATCAGGTGGAAGGACTCCATATCTAGCTCTTCGATATGATTTCAATTTGACTTAGATCTGT 2409
    |||||
DB 2137 GATCAGGTGGAAGGACTCCATATCTAGCTCTTCGATATGATTTCAATTTGACTTAGATCTGT 2196
QY 2410 GTGGGCATCCAGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAATGACAGAG 2469
```

```

Db      2197 GTGGGATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTAAATGACAGG 2256
Qy      2470 TCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGAT 2529
Db      2257 TCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGAT 2316
Qy      2530 ACAGGATACACGGAACGTTATATGAGGTACCTCTGACCAAGATGAACA 2576
Db      2317 ACAGGATACACGGAACGTTATATGAGGTACCTCTGACCAAGATGAACA 2363

RESULT 22
LOCUS   BE733691 1002 bp mRNA linear EST 15-SEP-2000
DEFINITION 60156288F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3844101 5',
mRNA sequence.
ACCESSION BE733691
VERSION BE733691.1 GI:10147786
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1002)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-femail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LHCMS39 row: p column: 22
High quality sequence stop: 672.
Location/Qualifiers
1. .1002
/organism="Homo sapiens"
/mol type="mRNA"
/db xref="taxon:9606"
/clone="IMAGE:3844101"
/lab host="DH10B (phage-resistant)"
/clone lib="NIH MGC 21"
/notes="Organ: placenta; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

```

Query Match 21.5%; Score 672; DB 7; Length 1002;
Best Local Similarity 100.0%; Fred. No. 0;
Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2226 CCCTTTGAATACCTAGCCTCTCTAGGTTATGTGTTGTAGTGATAGACACAGGGGATC 2285
Db      15 CCCTTTGAATACCTAGCCTCTCTAGGTTATGTGTTGTAGTGATAGACACAGGGGATC 74
Qy      2286 CTCTCCAGGCGCTTAATTTGAAGGGCCCTTTAAATATATAAATGGTCAATAGAAAT 2345
Db      75 CTGTCCAGGCGCTTAATTTGAAGGGCCCTTTAAATATATAAATGGTCAATAGAAAT 134
Qy      2346 TGACGATCAGGTGGGAAGGACTCCAATATCTAGCTTCTCGATPATGATTTCAATTCAGCTTAGA 2405
Db      135 TGACGATCAGGTGGGAAGGACTCCAATATCTAGCTTCTCGATPATGATTTCAATTCAGCTTAGA 194
```

```

Qy      2406 TCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTAAATGCA 2465
Db      195 TCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTAAATGCA 254
Qy      2466 GAGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTTA 2525
Db      255 GAGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTTA 314
Qy      2526 TGATACAGGATACACGGAACGTTATATGAGGTCACTCTGACCAAGATGAACAGGGCTATTTA 2585
Db      315 TGATACAGGATACACGGAACGTTATATGAGGTCACTCTGACCAAGATGAACAGGGCTATTTA 374
Qy      2586 CTTAGGATCTGTGCCATGCAAGCAAAAAGTTCCCTCTGAACCAAAATCGTTTACTGCT 2645
Db      375 CTTAGGATCTGTGCCATGCAAGCAAAAAGTTCCCTCTGAACCAAAATCGTTTACTGCT 434
Qy      2646 CTTACATGGTTTCTCTGGATGAGAAATGTCATTTTGCACATACCACTATATTACTGAGTTT 2705
Db      435 CTTACATGGTTTCTCTGGATGAGAAATGTCATTTTGCACATACCACTATATTACTGAGTTT 494
Qy      2706 TTTAGTGAGGCTGGAAAAGCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCAT 2765
Db      495 TTTAGTGAGGCTGGAAAAGCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCAT 554
Qy      2766 AAGAGTTCTCTGAATCGGAGAACATTTATGAATCTTTTGCACATACCTTTCAAGAAAA 2825
Db      555 AAGAGTTCTCTGAATCGGAGAACATTTATGAACTGTCATCTTTGCACTACCTTTCAAGAAAA 614
Qy      2826 CTTGGATCAGCTATTGCTGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 2885
Db      615 CTTGGATCAGCTATTGCTGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 674
Qy      2886 GGTATACACTGG 2897
Db      675 GGTATACACTGG 686

RESULT 23
LOCUS   CX787143 671 bp mRNA linear EST 02-MAR-2005
DEFINITION HSC3 81 D06.g1 A036 NIH MGC 260 Homo sapiens cDNA clone
IMAGE:7483429 5', mRNA sequence.
ACCESSION CX787143
VERSION CX787143.1 GI:58303933
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 671)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhardt, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-femail.nih.gov
Tissue Procurement: Bresagen, Inc.
cDNA Library Preparation: Express Genomics, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
University of Georgia
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM15806 row: g column: 11
Seq primer: JENREV (CAGGAACAGCTATGACC)
High quality sequence stop: 671.
Location/Qualifiers
1. .671
```

FEATURES

source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7483429"
/sex="male"
/tissue_type="human embryonic stem cells"
/cell_line="BG01"
/lab_host="PH108-T1 phage-resistant E. coli"
/notes="Vector: pExpress-1; Site 1: NotI; Site 2: EcoRV; RNA obtained from human embryonic stem cells isolated from the inner cell mass of blastocyst stage embryos. Cell line id and NIH Registry designation is BG01. Positive for SSEA3, SSEA4, Tra 1-81, CD9, Alk Phos, Oct4 and Nanog expression; negative for SSEA1 expression. Passage number 21. DNA primed using oligo-dr primer: 5'-pGACTAGTCTAGATCGGAGCGGCCCT(7)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below phred quality 16. Note: this is a Mammalian Gene Collection library."

```

ORIGIN

Query Match 21.5%; Score 671; DB 9; Length 671;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1650	TGTTGGGCTGCTCTCAAGTATTCAGTGTCCTATCAAGAGGAGATGACATATAC	1709
DB	1	TGTTGGGCTGCTCTCAAGTATTCAGTGTCCTATCAAGAGGAGATGACATATAC	60
QY	1710	CAGTGTGATGGGAAGTCTTGGCCGGATGGATCTAATATCAAGTTGATGAAGTCAG	1769
DB	61	CAGTGTGATGGGAAGTCTTGGCCGGATGGATCTAATATCAAGTTGATGAAGTCAG	120
QY	1770	AGCTGTGATATTTGAAGGACCAAGACTCCCTTTAGAGCATCACTGTACCTAGT	1829
DB	121	AGCTGTGATATTTGAAGGACCAAGACTCCCTTTAGAGCATCACTGTACCTAGT	180
QY	1830	CAGTACGTAATCTCGGAGAGGTGACCAAGGTGACTGACCGTGGTCTACATCTCTG	1889
DB	181	CAGTACGTAATCTCGGAGAGGTGACCAAGGTGACTGACCGTGGTCTACATCTCTG	240
QY	1890	CTGCATCAGTCAGCACTGTGACTCTTTTATAGTAAGTATAGTAACCAAGCAATCCCA	1949
DB	241	CTGCATCAGTCAGCACTGTGACTCTTTTATAGTAAGTATAGTAACCAAGCAATCCCA	300
QY	1950	CTGTGTGCTCTTCAAGCTATCAAGTCTCGAAGATGACCCCACTTGCACCAAGCA	2009
DB	301	CTGTGTGCTCTTCAAGCTATCAAGTCTCGAAGATGACCCCACTTGCACCAAGCA	360
QY	2010	ATTTGGGCGCACATTTTGGATTCAGCAGGTCTCTTCTGACTATCTCTCCAGAAAT	2069
DB	361	ATTTGGGCGCACATTTTGGATTCAGCAGGTCTCTTCTGACTATCTCTCCAGAAAT	420
QY	2070	TTTCTCTTTGAAAGTACTACTGATTTACATTTGATGGATGCTCTACAGCTCTATGA	2129
DB	421	TTTCTCTTTGAAAGTACTACTGATTTACATTTGATGGATGCTCTACAGCTCTATGA	480
QY	2130	TCTACAGCTCGAAGAAATATCTCTGCTGCTTCATATATGTTGCTCTCAGGTGCA	2189
DB	481	TCTACAGCTCGAAGAAATATCTCTGCTGCTTCATATATGTTGCTCTCAGGTGCA	540
QY	2190	GTGTGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCT	2249
DB	541	GTGTGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCT	600
QY	2250	AGGTTATGTGTTGTAGTATGATACCAAGGGGATCTGTCTCCAGAGGCTTAAATTTGA	2309
DB	601	AGGTTATGTGTTGTAGTATGATACCAAGGGGATCTGTCTCCAGAGGCTTAAATTTGA	660

QY 2310 AGCGGCTTTA 2320
 DB 661 AGCGGCTTTA 671

RESULT 24
 BU687687/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BU687687 804 bp mRNA linear EST 07-OCT-2002
 UI-CF-EC1-adu-o-02-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone
 UI-CF-EC1-adu-o-02-0-UI 3', mRNA sequence.
 BU687687
 EST.
 BU687687.1 GI:23543781
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 804)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 8889548
 Contact: McCray, PB
 McCray Lab
 University of Iowa
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
 Tel: 319 356 4866
 Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com) or from Open Biosystems
 (www.openbiosystems.com).
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

Source
 1. 804
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-CF-EC1-adu-o-02-0-UI"
 /tissue_type="Lung"
 /dev_stage="Adult and Fetal"
 /lab_hosts="DH10B (Life technologies) (T1 phage resistant)"
 /clone_lib="UI-CF-EC1"
 /note="Organ; Lung; Vector: pT7T3-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-CF-EC1 is a normalized cDNA library containing the
 following tissue(s): Normal lung from adult and from fetal
 day 64, day 87, week 19 and week 42. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dr primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pT7T3-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 ad118 tail. The sequence tag for this library is
 AAGTGTCTTAC.
 TAG ISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
 and 380-383
 TAG LIB=UI-CF-EC1
 TAG_SEQ=AAGTGTCTTAC"

ORIGIN

Query Match 21.2%; Score 663; DB 3; Length 804;

Db	549	TATGGGTCA	CGCTGACCT	CGATGAAG	AGGGGTATT	CTTAGGAT	CTGTGGCCAT	CGCAAGC	608
Qy	2610	AGAAAAGT	TTCCCTCT	CGAACCAAT	CGTTTAC	TGCTCTTAC	ATGTTTCT	CGATGAGAA	2659
Db	609	AGAAAAGT	TTCCCTCT	CGAACCAAT	CGTTTAC	TGCTCTTAC	ATGTTTCT	CGATGAGAA	668
Qy	2670	TGTCCATTT	TGCACAT	CACACCA	TATATAC	TACTGAG	TTTTT		2707
Db	669	TGTCATTTT	TGCATACC	ACCAAGT	ATATAC	TACTGAG	TTTTT		706

RESULT 29	
BC390103	
LOCUS	
DEFINITION	669 bp mRNA linear EST 12-MAR-2001
ACCESSION	602415944F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4524323 5',
VERSION	BC390103
KEYWORDS	BC390103.1 GI:13283649
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
	Hominidae; Homo.
REFERENCE	1. (bases 1 to 669)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

```

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Clone sequencing by: Incyte Genomics, Inc.
DNA distribution by: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L14M10428 row: g column: 12
High quality sequence stop: 669.
Location/Qualifiers
1. .669
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4524323"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC library."

```

	Query Match	20.6%	Score 644;	DB 2;	Length 669;
	Best Local Similarity	100.08;	Pred. NO. 0;		
	Matches 644;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
Qy	881	CTAACATCGTACACGAGAGAAAGGAGACTCACTTATGTGCACAAATGAGCTAGGCCACACA	940		
Db	1	CTAACATCGTAAACGAGAGAAAGGAGACTCACTTATGTGCACAAATGAGCTAGGCCACACA	60		
Qy	941	TGGAGAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTTCCAAGAGAGAAATTTGATA	1000		
Db	61	TGGAGAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTTCCAAGAGAGAAATTTGATA	120		
Qy	1001	GATATTCTGGCTATTGGTGGTGTCCAAAAGCTGCAACACTCCCAAGTGGTGGTAAAAATTC	1060		
Db	121	GATATTCTGGCTATTGGTGGTGTCCAAAAGCTGCAACACTCCCAAGTGGTGGTAAAAATTC	180		
Qy	1061	TTAGAAATCTTATATGAAGAAATATGATCATCTGAGTGGAGAAATATTATCATGTTTACATCC	1120		

Db	181	TTAGAAATTC	TATATGAAGAAAATG	AAATCTC	GAGTGGAAATTA	TCATGTTACATCCC	240	
QY	1121	CTATGTTGGAAA	CAAGGAGGCGAGATTC	ATTCCGTTTATCCT	TAACACAGGTACAGCAATC	1180		
Db	241	CTATGTTGGAAA	CAAGGAGGCGAGATTC	ATTCCGTTTATCCT	TAACACAGGTACAGCAATC	300		
QY	1181	CTAAAGTCAC	TTTTAAGATGTC	CAGAAATATGATTC	ATGATCGCTC	GAGGAAGGATCAT	PAGATG	1240
Db	301	CTAAAGTCAC	TTTTAAGATGTC	CAGAAATATGATTC	ATGATCGCTC	GAGGAAGGATCAT	PAGATG	360
QY	1241	TCATAGATAAG	GAAGAACTAAATCA	ACCTTTTGGAGATTC	TAATTTGAAGGAGT	TGAATATATG	1300	
Db	361	TCATAGATAAG	GAAGAACTAAATCA	ACCTTTTGGAGATTC	TAATTTGAAGGAGT	TGAATATATG	420	
QY	1301	CCAGAGCTGG	ATGGAATCTCTC	TGAGGGAATATGCTT	GGTCATCTACTAGATCGCTCCC	1360		
Db	421	CCAGAGCTGG	ATGGAATCTCTC	TGAGGGAATATGCTT	GGTCATCTACTAGATCGCTCCC	480		
QY	1361	AGACTCGCCT	CACAGATAGT	GTGTGATCTCA	CTCGAATTAATTTATCC	CAGTAGAAGATGATG	1420	
Db	481	AGACTCGCCT	CACAGATAGT	GTGTGATCTCA	CTCGAATTAATTTATCC	CAGTAGAAGATGATG	540	
QY	1421	TTATGGAAGG	CAGAGACTCATT	CGAGTCAGTGCCCTG	ATTCTGTGAGCCCACTAATATCT	1480		
Db	541	TTATGGAAGG	CAGAGACTCATT	CGAGTCAGTGCCCTG	ATTCTGTGAGCCCACTAATATCT	600		
QY	1481	ATGAAGAAC	ACACAGACATCTG	GAATAATATCC	ATGACATCTTT	1524		
Db	601	ATGAAGAAC	ACACAGACATCTG	GAATAATATCC	ATGACATCTTT	644		

FEATURES
source
High quality sequence stop: 669.
Location/Qualifiers
1. .669
/organism="Homo sapiens"
Accession
version
keywords
source
organism

REFERENCE AUTHORS	TITLE JOURNAL	COMMENT
1. J. H.
2. J. H.
3. J. H.
4. J. H.
5. J. H.
6. J. H.
7. J. H.
8. J. H.
9. J. H.
10. J. H.
11. J. H.
12. J. H.
13. J. H.
14. J. H.
15. J. H.
16. J. H.
17. J. H.
18. J. H.
19. J. H.
20. J. H.
21. J. H.
22. J. H.
23. J. H.
24. J. H.
25. J. H.
26. J. H.
27. J. H.
28. J. H.
29. J. H.
30. J. H.
31. J. H.
32. J. H.
33. J. H.
34. J. H.
35. J. H.
36. J. H.
37. J. H.
38. J. H.
39. J. H.
40. J. H.
41. J. H.
42. J. H.
43. J. H.
44. J. H.
45. J. H.
46. J. H.
47. J. H.
48. J. H.
49. J. H.
50. J. H.
51. J. H.
52. J. H.
53. J. H.
54. J. H.
55. J. H.
56. J. H.
57. J. H.
58. J. H.
59. J. H.
60. J. H.
61. J. H.
62. J. H.
63. J. H.
64. J. H.
65. J. H.
66. J. H.
67. J. H.
68. J. H.
69. J. H.
70. J. H.
71. J. H.
72. J. H.
73. J. H.
74. J. H.
75. J. H.
76. J. H.
77. J. H.
78. J. H.
79. J. H.
80. J. H.
81. J. H.
82. J. H.
83. J. H.
84. J. H.
85. J. H.
86. J. H.
87. J. H.
88. J. H.
89. J. H.
90. J. H.
91. J. H.
92. J. H.
93. J. H.
94. J. H.
95. J. H.
96. J. H.
97. J. H.
98. J. H.
99. J. H.
100. J. H.

TITLE	JOURNAL	COMMENT

KZFD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: info@kzfd.de

E-mail: www.rzpd.de
RZPD; RZPDp9016P133.

RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
<http://www.rzpd.org/>

www.rzpu.de/cgi-bin/products/set.cgi?libNo=9016
Inge Arlart

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6 D-14105 Berlin

Tel: +49 30 32639 100
 www.burwey.de, D-14059 Berlin, Germany

Fax: +49 30 32639 111
www.rzpd.de

This clone is available from RZPD;

http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=RZPDP99

Primer name: qe3_4 , Primer sequence: CCGATAACCAATTCACACAC

Location/Qualifiers
1. 887

```
1. .001
   /organism="Homo sapiens"
```

```

/mol_type="mRNA"
/db_xref="taxon:9606"

```

```
/w_xref="taxon:9606"
/clone="RZPDp9016P133"
```

Db	781	ATACGGATGGATCC	794	
RESULT 31				
BU622228/c				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
FEATURES				
source				
ORIGIN				
Query Match				
Best Local Similarity				
Matches 791; Conservative				
0; Mismatches				
3; Indels				
0; Gaps				
QY	25	GGCGCGTGTACTGCGCGCTGCTTCTTAGTGGCGGCTTGGCGGCTGGTGTTCACCG	84	
Db	1	GGCGCGTGTACTGCGCGCTGCTTCTTAGTGGCGGCTTGGCGGCTGGTGTTCACCG	60	
QY	85	GGCGCGCGCGCGGAGGACCACTGCAACAGGACCGGAGTGGAGCGGCGCAGCATGAAG	144	
Db	61	GGCGCGCGCGCGGAGGACCACTGCAACAGGACCGGAGTGGAGCGGCGCAGCATGAAG	120	
QY	145	CGCGCGAGCGCGCTCCATAGCGCACTGTCGGGACGCTCGCGCGGCGGCGGCGGGAAGCA	204	
Db	121	CGCGCGAGCGCGCTCCATAGCGCACTGTCGGGACGCTCGCGCGGCGGCGGCGGGAAGCA	180	
QY	205	AAATGCACATGCGACGACCAATGGAACAGAAACAGCTGGGTGTGAGATATTTGAAACT	264	
Db	181	AAATGCACATGCGACGACCAATGGAACAGAAACAGCTGGGTGTGAGATATTTGAAACT	240	
QY	265	GGCGACTGTGAGGAGAAATATGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTT	324	
Db	241	GGCGACTGTGAGGAGAAATATGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTT	300	
QY	325	GAGCGTATTCCTGGAGTCACTTAAAGCTGCTTGGCGGATACACAGAAATATCATGGC	384	
Db	301	GAGCGTATTCCTGGAGTCACTTAAAGCTGCTTGGCGGATACACAGAAATATCATGGC	360	
QY	385	TACATGATGCTTAAGGACCAACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGA	444	
Db	361	TACATGATGCTTAAGGACCAACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGA	420	
QY	445	CCTCATTCAGACAGATCTTATACCTTGGCATCTGCTGGAGACAGACAGAAATATCAGCTG	504	
Db	421	CCTCATTCAGACAGATCTTATACCTTGGCATCTGCTGGAGACAGACAGAAATATCAGCTG	480	
QY	505	TTTTTATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAA	564	
Db	481	TTTTTATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAA	540	
QY	565	CCTCTTTTGGATCTTTTTCAGGCAACATCTGACATGATGATGATGATGATGATGATGATGAT	624	
Db	541	CCTCTTTTGGATCTTTTTCAGGCAACATCTGACATGATGATGATGATGATGATGATGATGAT	600	
QY	625	CTATTAGAGAAAGAAACCGATTCGAAACAGCTGGAAATTCGTTTCTTACGATTTATCACC	684	
Db	601	CTATTAGAGAAAGAAACCGATTCGAAACAGCTGGAAATTCGTTTCTTACGATTTATCACC	660	
QY	685	GGAAGTGGAAACATTTCTGTTTCAAGCGCGGTAGTGGAAATTTATCAGTTAAAGATGGAGG	744	
Db	661	GGAAGTGGAAACATTTCTGTTTCAAGCGCGGTANTGCAATTTATCAGTTAAAGATGGAGG	720	
QY	745	CCACAGGATTTAGCCACACACCTTTAAGGCCCAATCTAGTGGAACTAGTGTGCCAAC	804	
Db	721	CCACAGGATTTAGCCACACACCTTTAAGGCCCAATCTAGTGGAACTAGTGTGCCAAC	780	
QY	805	ATACGGATGGATCC	818	


```

Qy 3084 GACCCAGCAATACCATTAAGATTACTAAAAA 3120
Db 44 GACCCAGCAATACCATTAAGATTACTAAAAA 8

RESULT 33
BI223892 853 bp mRNA linear EST 11-JUL-2001
LOCUS 602941035F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5104113 5',
DEFINITION mRNA sequence.
ACCESSION BI223892
VERSION BI223892.1 GI:14677336
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 853)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1251 row: e column: 10
High quality sequence stop: 718.
FEATURES
source
1..853
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5104113"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH MGC 12"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN
Query Match 20.1%; Score 627; DB 2; Length 853;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 677; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1527 TGTTCCTCCCAAGTCAGAGAGAAATTGAGTTATTTTCCTCTGAATGCAAAAC 1586
Db 43 TGTTCCTCCCAAGTCAGAGAGAAATTGAGTTATTTTCCTCTGAATGCAAAAC 102

Qy 1587 AGTTTCCTGTCATTTATACAAATACATCTATTTTAAAGGAAAGCAAAATATAAAGCATC 1646
Db 103 AGTTTCCTGTCATTTATACAAATACATCTATTTTAAAGGAAAGCAAAATATAAAGCATC 162

Qy 1647 CAGTGTGGCTGCTGCTCCAGTGATTTTCAAGTGCTCTATCAAGAGGATAGCAAT 1706
Db 163 CAGTGTGGCTGCTGCTCCAGTGATTTTCAAGTGCTCTATCAAGAGGATAGCAAT 222

Qy 1707 TACCAGTGTGAATGGAAAGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGAAGT 1766
Db 223 TACCAGTGTGAATGGAAAGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGAAGT 282

Qy 1767 CAGAGGCTGGTATATTTTGAAGGCAACCAAGATCCCTTTAGAGATCACCTGTACGT 1826
Db 283 CAGAGGCTGGTATATTTTGAAGGCAACCAAGATCCCTTTAGAGATCACCTGTACGT 342

Qy 1827 AGTCAGTTACGTAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTC 1886

```


QY 608 ATTCTGAGAGAACTATTATAGAGAAAGAAACGATTTGGACAGTCGGAAATGCTTT 667
|||||
Db 607 ATTCTGAGAGAGAGACTATTATAGAGAAAGAAACGATTTGGACAGTCGGAAATGCTTT 666
|||||
QY 668 CTTACGATTATCACCAA 684
|||||
Db 667 CTTACGATTATCACCAA 683
|||||

RESULT 36
BG715866
LOCUS 602675282F1 NIH_MGC_96 695 bp mRNA linear EST 08-MAY-2001
DEFINITION 602675282F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4797669 5',
mRNA sequence.
ACCESSION BG715866
VERSION BG715866.1 GI:13995053
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 695)
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10684 row: d column: 22
High quality sequence stop: 695.
Location/Qualifiers
1. .695
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4797669"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gtcgag); Oligo-dT primed primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to R0T 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN
Query Match 20.0%; Score 623; DB 2; Length 695;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 673; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2025 TTGGATTTCAGCAGGTCTCTTCCTGACTATCTCCAGAAATTTCTTTTCAAG 2084
|||||
Db 22 TTGGATTTCAGCAGGTCTCTTCCTGACTATCTCCAGAAATTTCTTTTCAAG 81
|||||

QY 2085 TACTACTGATTATACATGATGGGATGCTCTACAAGCCTCATGATCTACAGCTCGAAA 2144
|||||
Db 82 TACTACTGATTATACATGATGGGATGCTCTACAAGCCTCATGATCTACAGCTCGAAA 141
|||||

QY 2145 GAAATATCCTACTGTGCTGTTTCATATATGGTGGTCCCTCAGTGCGAGTTGGTGAATAATCG 2204
|||||

Db 142 GAAATATCCTACTGTGCTGTTTCATATATGGTGGTCCCTCAGTGCGAGTTGGTGAATAATCG 201
|||||

QY 2205 GTTTAAAGAGAGTCAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGGTTGT 2264
|||||

Db 202 GTTTAAAGAGAGTCAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGGTTGT 261
|||||

QY 2265 AGTGATAGACAAACAGGGGATCCTGTCCACCGAGGGCTTAAATTTTGAAGGCGCTTTAAATA 2324
|||||

Db 262 AGTGATAGACAAACAGGGGATCCTGTCCACCGAGGGCTTAAATTTTGAAGGCGCTTTAAATA 321
|||||

QY 2325 TAAATGGGTCAATTAAGAAATTTGACGNTCAGSTGGAAGGACTCCAAATATCTAGTCTTCG 2384
|||||

Db 322 TAAATGGGTCAATTAAGAAATTTGACGNTCAGSTGGAAGGACTCCAAATATCTAGTCTTCG 381
|||||

QY 2385 ATATGATTTTCATTGACTTTAGATCGTGTGGGCATCCACGGCTGGTCTATATGGAGGATACCT 2444
|||||

Db 382 ATATGATTTTCATTGACTTTAGATCGTGTGGGCATCCACGGCTGGTCTATATGGAGGATACCT 441
|||||

QY 2445 CTCCTGATGGCATTAAATGACGAGCTCAGATATCTTCAGGGTTCCTATTTGCTGGGGCCCC 2504
|||||

Db 442 CTCCTGATGGCATTAAATGACGAGCTCAGATATCTTCAGGGTTCCTATTTGCTGGGGCCCC 501
|||||

QY 2505 AGTCACCTGTGTGGATCTTCTATGATACAGGATACACGGAACTTTATATGGTTCACCCCTGA 2564
|||||

Db 502 AGTCACCTGTGTGGATCTTCTATGATACAGGATACACGGAACTTTATATGGTTCACCCCTGA 561
|||||

QY 2565 CCAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTC 2624
|||||

Db 562 CCTGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTC 621
|||||

QY 2625 TGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAAATGTCATTTTGCACA 2684
|||||

Db 622 TGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAAATGTCATTTTGCACA 681
|||||

QY 2685 TACCAATATATAC 2698
|||||

Db 682 TACCAATATATAC 695
|||||

RESULT 37
AI819365/c
LOCUS AI819365 634 bp mRNA linear EST 07-MAR-2000
DEFINITION wg61c02.x1 Soares NSF P8_9W_OT_PA_P.S1 Homo sapiens cDNA clone
IMAGE:2369570 3' similar to TR:075273 O75273 R26984_1 ; mRNA
sequence.
ACCESSION AI819365
VERSION AI819365.1 GI:5438444
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 634)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 699 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 469.
Location/Qualifiers
1. .634
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2369570"
/lab_host="DH10B"

FEATURES
source

```
/clone lib="Soares NSF P8 9W OT PA_P_S1"
/notes="Organ: pooled; Vector: p773D-Paci; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from five
normalized libraries were mixed, and ss circles were made
in vitro. Following HAP purification, this DNA was used as
tracer in a substructure hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 5 libraries. The pools consisted of the
following libraries and clones: Soares NBHSF pool 1:
309384-310919, 323208-325895 Soares NB2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares NB2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NBHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NBHOT
pool 1: 723770-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."
```

ORIGIN

```
Query Match 19.7%; Score 614; DB 1; Length 634;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2499 GGGCCAGTCACTCTGGATCTCTATGATACAGATACACGGAACGTATATGGGTCA 2558
DB 614 GGGCCAGTCACTCTGGATCTCTATGATACAGATACACGGAACGTATATGGGTCA 2558

QY 2559 CCCTGACGAGATGACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTT 2618
DB 554 CCCTGACGAGATGACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTT 2618

QY 2619 CCCTCTGACCAAAATGTTTACTGCTCTTACATGTTTCTCGATGAGATGCAATTT 2678
DB 494 CCCTCTGACCAAAATGTTTACTGCTCTTACATGTTTCTCGATGAGATGCAATTT 2678

QY 2679 TGCATACAGATATTACTCAAGTCTTTAGTAGGGCTGGAAGCCATATGATTTACA 2738
DB 434 TGCATACAGATATTACTCAAGTCTTTAGTAGGGCTGGAAGCCATATGATTTACA 2738

QY 2739 GATCTATCTCAGGAGACACAGATAGAGTTCTCGAATCGGAGACATATGAACT 2798
DB 374 GATCTATCTCAGGAGACACAGATAGAGTTCTCGAATCGGAGACATATGAACT 2798

QY 2799 GCATCTTTGCACTACCTTCAAGAAACCTTGGATACGATTTGCTGCTCTAAAGTGT 2858
DB 314 GCATCTTTGCACTACCTTCAAGAAACCTTGGATACGATTTGCTGCTCTAAAGTGT 2858

QY 2859 ATAAATTTGACTGTAGAACTCTCTGGTATACACTTGGCTATTTAAACCAATGAGGAG 2918
DB 254 ATAAATTTGACTGTAGAACTCTCTGGTATACACTTGGCTATTTAAACCAATGAGGAG 2918

QY 2919 TTTAATCAACAGAAACACAGAAATGATCATCATTTTGTATACCTGCTATTTAACCNAATGGGAGG 195
DB 194 TTTAATCAACAGAAACACAGAAATGATCATCATTTTGTATACCTGCTATTTAACCNAATGGGAGG 195

QY 2979 ACTCTGAAATAAATGTGGTCCATGCAAGGCTCTACGGTTTGTGTAGTAATCTAATA 3038
DB 134 ACTCTGAAATAAATGTGGTCCATGCAAGGCTCTACGGTTTGTGTAGTAATCTAATA 3038

QY 3039 CTTTAACCCACATGCTCAAAATCAATATGATATATTTCTTGAGAGACCCAGCAATACCA 3098
DB 74 CTTTAACCCACATGCTCAAAATCAATATGATATATTTCTTGAGAGACCCAGCAATACCA 3098

QY 3099 TAAGAATTAATAA 3112
DB 14 TAAGAATTAATAA 3112
```

```
RESULT 38
BQ675006
LOCUS
DEFINITION
BQ675006
5', mRNA sequence.
AGENCY: 8303592 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6274622
EST 15-JUL-2002
```

VERSION
KEYWORDS
SOURCE
ORGANISM

BQ675006.1 GI:21785840
EST.
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 957)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2456 row: h column: 15
High quality sequence start: 4
High quality sequence stop: 624.
Location/Qualifiers
1..957

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6274622"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_102"
/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI. cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."

ORIGIN

```
Query Match 19.5%; Score 607; DB 3; Length 957;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 607; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1633 AAATATAACCATCCAGTGGCTGGCTGCTCCCAAGTGATTTCAAGTGCTCTATCAA 1692
DB 12 AAATATAACCATCCAGTGGCTGGCTGCTCCCAAGTGATTTCAAGTGCTCTATCAA 1692

QY 1693 GAGGAGATGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGCGATGGAATCTAATATC 1752
DB 72 GAGGAGATGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGCGATGGAATCTAATATC 1752

QY 1753 CAAGTTGATGAGTCAAGAGGCTGGTGAATTTTGAAGGCACCAAGAGCTCCCGCTTTAGAG 1812
DB 132 CAAGTTGATGAGTCAAGAGGCTGGTGAATTTTGAAGGCACCAAGAGCTCCCGCTTTAGAG 1812

QY 1813 CATCACTGTACAGTCAAGTGTACGTAATCTGAGAGGTCACAAAGCTGACGCTGACCGT 1872
DB 192 CATCACTGTACAGTCAAGTGTACGTAATCTGAGAGGTCACAAAGCTGACGCTGACCGT 1872

QY 1873 GGCTACTCATCTTTCTGTCATCAGTCAAGTGTGACTTCTTTTATAAGTAAGTATAGT 251
DB 252 GGCTACTCATCTTTCTGTCATCAGTCAAGTGTGACTTCTTTTATAAGTAAGTATAGT 251

QY 1933 AACAGAGAAATCCACACTGTGTGTCCCTTTTACAGCTATCAAGCTGCTGAAGTACCCA 311
DB 312 AACAGAGAAATCCACACTGTGTGTCCCTTTTACAGCTATCAAGCTGCTGAAGTACCCA 311

QY 1993 ACTTGCAAAACAAAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGTCTCTCTCTGAC 2052
DB 372 ACTTGCAAAACAAAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGTCTCTCTCTGAC 431
```


Thu Jun 22 09:04:26 2006

ORIGIN

Query Match 19.2%; Score 598; DB 8; Length 612;
 Best Local Similarity 100.0%; Pred. No. 1.1e-311;
 Matches 598; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 201 AGGAAATGCAACATGGCAGCAGCAATGGAACAGCAAGCTGGGTGTGGAGATATTTGA 260
 DB 15 AGGAAATGCAACATGGCAGCAGCAATGGAACAGCAAGCTGGGTGTGGAGATATTTGA 74
 QY 261 AACTGCGGAGCTGTGAGGAGAAATTAATCAAGATCGGCTTAATTTGGAGCCCTTTTA 320
 DB 75 AACTGCGGAGCTGTGAGGAGAAATTAATCAAGATCGGCTTAATTTGGAGCCCTTTTA 134
 QY 321 TGTGAGCGGTATTTCTGGAGTCAGCTTAAAGAGCTTGGCCGATACAGAGAAATATCA 380
 DB 135 TGTGAGCGGTATTTCTGGAGTCAGCTTAAAGAGCTTGGCCGATACAGAGAAATATCA 194
 QY 381 TGGCTACATGATGCTTAAGCAGCACATGATTTTCAATGTTGTGAAGAGAAATGATCCAGA 440
 DB 195 TGGCTACATGATGCTTAAGCAGCACATGATTTTCAATGTTGTGAAGAGAAATGATCCAGA 254
 QY 441 TGGACCTCATTTACAGAGAAATTTTACCTTGGCATGCTGTGGTGAAGAGAAATATCA 500
 DB 255 TGGACCTCATTTACAGAGAAATTTTACCTTGGCATGCTGTGGTGAAGAGAAATATCA 314
 QY 501 ACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGACAGCAGCTTAAATGCTCTCTTG 560
 DB 315 ACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGACAGCAGCTTAAATGCTCTCTTG 374
 QY 561 GAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGAGTATGGAATGATTTCTCGAGAAGA 620
 DB 375 GAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGAGTATGGAATGATTTCTCGAGAAGA 434
 QY 621 AGAATCTATTAAGAGAAAGAAACGATTTGGAACAGTGGAAATGCTTCTTACGATATCA 680
 DB 435 AGAATCTATTAAGAGAAAGAAACGATTTGGAACAGTGGAAATGCTTCTTACGATATCA 494
 QY 681 CCAAGGAGTGGAAATTTCTGTTTCAAGCGGTAGTGGAAATTTATCACTTAAAGATGG 740
 DB 495 CCAAGGAGTGGAAATTTCTGTTTCAAGCGGTAGTGGAAATTTATCACTTAAAGATGG 554
 QY 741 AGGCGCAACAGGATTTACGCAACACTTTTAAAGCCCAATCTAGTGAAGAACTAGTTGT 798
 DB 555 AGGCGCAACAGGATTTACGCAACACTTTTAAAGCCCAATCTAGTGAAGAACTAGTTGT 612

RESULT 40

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CB153590 596 bp mRNA linear EST 29-JAN-2003
 K-EST0211080 B1T694954 Homo sapiens cDNA clone B1T694954-33-D08 5',
 mRNA sequence.
 CB153590
 CB153590.1 GI:28138584
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 1 (bases 1 to 596)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongseung@mail.kribb.re.kr

QY 2053 TATACCTCCAGAAATTTCTTTTGAAGTACTACTGATTTACATTTGATGGATG 2112
 DB 432 TATACCTCCAGAAATTTCTTTTGAAGTACTACTGATTTACATTTGATGGATG 491
 QY 2113 CTCTACAGGCTCATGATCTACAGCCTGGAAGAAATATCTCTAGTGTGTTTCATATAT 2172
 DB 492 CTCTACAGGCTCATGATCTACAGCCTGGAAGAAATATCTCTAGTGTGTTTCATATAT 551
 QY 2173 GGTGGTCTCAGGTGCGAGTTGGTGAATATCGTTTAAAGAGAGTCAAGATTTCCCGTTG 2232
 DB 552 GGTGGTCTCAGGTGCGAGTTGGTGAATATCGTTTAAAGAGAGTCAAGATTTCCCGTTG 611
 QY 2233 AATACCC 2239
 DB 612 AATACCC 618

RESULT 39

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CV024163 612 bp mRNA linear EST 20-AUG-2004
 1359 Full Length cDNA from the Mammalian Gene Collection Homo
 sapiens cDNA 5' similar to BC030688, mRNA sequence.
 CV024163
 CV024163.1 GI:51481992
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 1 (bases 1 to 612)
 Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S.,
 Driscot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O.,
 Clingingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T.,
 Simmons,B., Sequerra,R., Boeak,S., Doucette-Stamm,L., Le Peuch,C.,
 Vandenhaute,J., Cubick,M.E., Albala,J.S., Hill,D.E. and Vidal,M.
 Human ORFome Version 1.1: a Platform for Reverse Proteomics
 Genome Res. (2004) In press
 Contact: Vidal M
 Marc Vidal Laboratory
 Dana Farber Cancer Institute
 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
 Tel: 617 632 5180
 Fax: 617 632 5739
 Email: Marc.Vidal@dfci.harvard.edu
 ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF
 results from a PCR reaction using an MGC full-length cDNA as
 template DNA and ORF specific primers
 PCR Primers
 FORWARD: ATGAATAAATACAGAAATGCAACATGG
 BACKWARD: TATATCACTTTTATAGCAGCAATAC
 Insert Length: 612 Std Error: 60.00
 Plate: 11079 row: 01 column: D
 Seq primer: ACTGGCGTCTGTTTACAAACGTCGTGACTGGGAAAC
 High quality sequence start: 103
 High quality sequence stop: 611
 POLYA=NO.

FEATURES
 source
 Location/Qualifiers
 1..612
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="mixed"
 /clone_lib="Full length cDNA from the Mammalian Gene
 Collection"
 /notes="Vector: mixed; The ORFs were PCR amplified from the
 MGC (Mammalian Gene Collection) as of April 2004 and
 cloned by recombinational Gateway cloning into pDONR223
 Donor vector. Reference : MGC (Mammalian Gene Collection)
 Program Team, Generation and Initial Analysis of more than
 15,000 Full-Length Human and Mouse cDNA Sequences. PNAS,
 2002, 99(26), 16899-16903"

QY	54
Db	1
QY	114
Db	61
QY	174
Db	121
QY	234
Db	181
QY	294
Db	241

Thu Jun 22 09:04:26 2006

	Matches	587; Conservative	0; Mismatches	0; Indels	0; Gaps	0;
QY	1	AAGTGTCTAAAGCCCTTCGAGGCCAAGCGCGCTACTCCTCCCGCGCTGCTTTTAGTGCCG	60			
Db	1	AAGTGTCTAAAGCCCTTCGAGGCCAAGCGCGCTACTCCTCCCGCGCTGCTTTTAGTGCCG	60			
QY	61	GTTTGGCGCGCTGGTGTCAACCGGCGCGCGCGCGAGAGAGCCTCACTGAACACGACCG	120			
Db	61	GTTTGGCGCGCTGGTGTCAACCGGCGCGCGCGCGAGAGAGCCTCACTGAACACGACCG	120			
QY	121	GAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGCACGTCGCGGACGG	180			
Db	121	GAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGCACGTCGCGGACGG	180			
QY	181	TGCGGCGCGCGCGCGGAGGAAAAATGCAATGCGAGCAGCATGAAAACAACAG	240			
Db	181	TGCGGCGCGCGCGCGGAGGAAAAATGCAATGCGAGCAGCATGAAAACAACAG	240			
QY	241	CTGGGTGTTGAGATATTTGAAACTGCGGAGCTGTGAGGAGATATTGAATCACAGGATCG	300			
Db	241	CTGGGTGTTGAGATATTTGAAACTGCGGAGCTGTGAGGAGATATTGAATCACAGGATCG	300			
QY	301	CCTAAATGAGCGCTTTTTTATGTTGAGCGGTATTCTTGGAGTCACTTAATAAAGCTGCTT	360			
Db	301	CCTAAATGAGCGCTTTTTTATGTTGAGCGGTATTCTTGGAGTCACTTAATAAAGCTGCTT	360			
QY	361	GCGGATACCAAGAAAAATATCATGGCTACATGATGCGTAAAGCACCACATGATTTTCATGTTT	420			
Db	361	GCGGATACCAAGAAAAATATCATGGCTACATGATGCGTAAAGCACCACATGATTTTCATGTTT	420			
QY	421	GTGAAGAGGAGTATCCAGATGACCTCAATCAGACAGAATCTATTACCTTGCATGCTCT	480			
Db	421	GTGAAGAGGAGTATCCAGATGACCTCAATCAGACAGAATCTATTACCTTGCATGCTCT	480			
QY	481	GSTGAGAACACAGAAAAATACACTGTTTATCTGAAATCTCCCAAATCTCAATAGAGCA	540			
Db	481	GSTGAGAACACAGAAAAATACACTGTTTATCTGAAATCTCCCAAATCTCAATAGAGCA	540			
QY	541	GCAGTCTTAATGCTCTCTTGGAGCCCTCTTTTGGATCTTTTTCAGGC	587			
Db	541	GCAGTCTTAATGCTCTCTTGGAGCCCTCTTTTGGATCTTTTTCAGGC	587			
RESULT 43						
LOCUS	DA441156	586 bp	mRNA	linear	EST 05-NOV-2005	
DEFINITION	DA441156 CTONG2 Homo sapiens cDNA clone CTONG2010006 5', mRNA sequence.					
ACCESSION	DA441156					
VERSION	DA441156.1	GI:80838779				
KEYWORDS	EST.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	1 (bases 1 to 586) Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaquri,H., Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushiida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.					
TITLE	Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes					
JOURNAL	PUBMED	Genome Res. 16 (1), 55-65 (2006)				
COMMENT	Contact: Takao Isogai FLJ Project (HRI Team) Helix Research Institute 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: flj-cdnafinity.com NEO human cDNA project (New Energy and Industrial Technology Development Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB. Location/Qualifiers 1..587 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="BRTHA2030154" /tissue_type="chalamus" /clone_lib="BRTHA2" /note="Vector: pME183FL3"					
FEATURES	source					
ORIGIN						

Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEBO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan): cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
Location/Qualifiers
1. .596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CTONG2010006"
/tissue_type="tongue, tumor tissue"
/clone_lib="CTONG2"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 18.8%; Score 586; DB 9; Length 586;
Best Local Similarity 100.0%; Pred. No. 3.6e-305;
Matches 586; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 AAAGCCTCCGAGGCCAAGCGCTGCTACTGCGCGCGCTGCTTCTTAGTCCGGTTCGC 67
Db 1 AAAGCCTCCGAGGCCAAGCGCTGCTACTGCGCGCGCTGCTTCTTAGTCCGGTTCGC 60
QY 68 CGCTTGGTGTGTACCG 127
Db 61 CGCTTGGTGTGTACCG 120
QY 128 GCGCGCGCAGCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 187
Db 121 GCGCGCGCAGCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 188 GCG 247
Db 181 GCG 240
QY 248 TTGAGATATTTGAACTCGCGGCTGAGGAGATATTTGAATCAGAGGATCGGCTTAAT 307
Db 241 TTGAGATATTTGAACTCGCGGCTGAGGAGATATTTGAATCAGAGGATCGGCTTAAT 300
QY 308 TGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGTTGCGGATA 367
Db 301 TGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGTTGCGGATA 360
QY 368 CCAGAAATATCATGGCTACATGATGCTAAGGCACCATGATTTTCAITTTTGTGAAGA 427
Db 361 CCAGAAATATCATGGCTACATGATGCTAAGGCACCATGATTTTCAITTTTGTGAAGA 420
QY 428 GGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTCCCATGTCTGTGAGA 487
Db 421 GGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTCCCATGTCTGTGAGA 480
QY 488 ACAGAGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGACAGCACTCT 547
Db 481 ACAGAGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGACAGCACTCT 540
QY 548 TAATGCTCTTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACT 593
Db 541 TAATGCTCTTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACT 586

RESULT 44
DB157274
LOCUS
DEFINITION
DB157274 THYMUS3 Homo sapiens cDNA clone THYMUS3036274 5', mRNA
sequence.
DB157274
VERSION
DB157274.1 GI:83205040
KEYWORDS
Homo sapiens (human)
EST. 04-DEC-2005

ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Makaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi, F., Ii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T., and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEBO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan): cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
Location/Qualifiers
1. .593
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="THYMUS3036274"
/tissue_type="thymus"
/clone_lib="THYMUS3"
/note="Vector: pME18SFL3"

FEATURES
source

Query Match 18.8%; Score 586; DB 9; Length 593;
Best Local Similarity 100.0%; Pred. No. 3.6e-305;
Matches 586; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGTCTAAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTCCGC 60
Db 8 AAGTCTAAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTCCGC 67
QY 61 CGTTTCGCGCGCTGGTGTGTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 68 CGTTTCGCGCGCTGGTGTGTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 127
QY 121 GAGTGGAGCG 180
Db 128 GAGTGGAGCG 187
QY 181 TCCG 240
Db 188 TCCG 247
QY 241 CTGGGTGTTGAGATATTTGAACTCGCGGATCTGAGGAGATATTTGAATCAGAGGATCGG 300
Db 248 CTGGGTGTTGAGATATTTGAACTCGCGGATCTGAGGAGATATTTGAATCAGAGGATCGG 307
QY 301 CTTAAATTTGAGCGCTTTTATTTGAGCGGTATTTCTGGAGTCAGCTTAAAGAGCTGCTT 360
Db 308 CTTAAATTTGAGCGCTTTTATTTGAGCGGTATTTCTGGAGTCAGCTTAAAGAGCTGCTT 367
QY 361 GCGGATACAGAAATATCATGCTACATGATGGCTAAGGCGCGCGCGCGCGCGCGCGCG 420
Db 368 GCGGATACAGAAATATCATGCTACATGATGGCTAAGGCGCGCGCGCGCGCGCGCGCG 427
QY 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTCCCATGCT 480

ORIGIN

Thu Jun 22 09:04:26 2006

Db	428	GTGAAGAGGAATGATCCAGATGACCTATTTCAGACAGAACTATTACCTTGGCATGCT	487
Qy	481	GGTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGAGCA	540
Db	488	GGTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGAGCA	547
Qy	541	GCAGTCTTAATGCTCTCTTGGAGCCCTTTTGGATCTTTTCAGG	586
Db	548	GCAGTCTTAATGCTCTCTTGGAGCCCTTTTGGATCTTTTCAGG	593
RESULT 45			
DB159782			
LOCUS	DB159782	THYM3 Homo sapiens cDNA clone	THYM3039682 5', mRNA
DEFINITION			sequence. EST 04-DEC-2005
ACCESSION	DB159782		
VERSION	DB159782.1	GI:83253907	
KEYWORDS			EST.
SOURCE			Homo sapiens (human)
ORGANISM			Homo sapiens
REFERENCE			1 (bases 1 to 587)
AUTHORS			Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, I., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
TITLE			Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
JOURNAL			Genome Res. 16 (1), 55-65 (2006)
PUBMED			16344560
COMMENT			Contact: Takao Isogai FLJ Project (HRI Team) Helix Research Institute 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: flj-cdna@nifty.com NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.
FEATURES			source
ORIGIN			
Query Match			18.8%; Score 585; DB 9; Length 587;
Best Local Similarity			100.0%; Pred. NO. 1.3e-304;
Matches			585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	AAGTGCTTAAGCTCCGAGGCCAAGCGCTGCTACTGCGCGCTTCTTAGTCCG	60
Db	3	AAGTGCTTAAGCTCCGAGGCCAAGCGCTGCTACTGCGCGCTTCTTAGTCCG	62
Qy	61	GCTTGGCGCTCGGCTTGTTCACCGGCGCGCGCGAGGAGCCACTCAACCGAGCCG	120
Db	63	GCTTGGCGCTCGGCTTGTTCACCGGCGCGCGCGAGGAGCCACTCAACCGAGCCG	122

Db	541	GGACTATGGAATGTATTCTCGAGAAGAACTATTAAAGAGA	582
RESULT 48			
BX645431			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
QY	730	GTAAAGATGAGGGCCCAAGGATTTAGCAACACCTTTAAGGCCCAATCTAGTGGAA	789
Db	1	GTAAAGATGAGGGCCCAAGGATTTAGCAACACCTTTAAGGCCCAATCTAGTGGAA	60
QY	790	ACTAGTTGTCTCCAAACATACCGATGGATCCAAATTTATGCGCGCTGATCCAGATGGATT	849
Db	61	ACTAGTTGTCTCCAAACATACCGATGGATCCAAATTTATGCGCGCTGATCCAGATGGATT	120
QY	850	GCTTTTATACATCAACGATATTTGGATATCTAACATCTAACACGAGAGAAAGGAGA	909
Db	121	GCTTTTATACATCAACGATATTTGGATATCTAACATCTAACACGAGAGAAAGGAGA	180
QY	910	CTCATTTATGTGCAATAGTAGTCCCAACATGGAAGAGATCCAGATTCAGCTGGAGTC	969
Db	181	CTCATTTATGTGCAATAGTAGTCCCAACATGGAAGAGATCCAGATTCAGCTGGAGTC	240
QY	970	GCTACCTTTGTCTCCAAAGAGATTTGATAGATTTCTGGCTATTTGGTGGTGTCCAAA	1029
Db	241	GCTACCTTTGTCTCCAAAGAGATTTGATAGATTTCTGGCTATTTGGTGGTGTCCAAA	300
QY	1030	GCTGAACAACTCCCACTGGTGGTAAAAATTTCTAGAAATTTCTATATGAAGAAATGATGAA	1089
Db	301	GCTGAACAACTCCCACTGGTGGTAAAAATTTCTAGAAATTTCTATATGAAGAAATGATGAA	360
QY	1090	TCTGAGGTGGAAATTTATTCATGTTTACATCCCTTATCTGTTGAAAACAAGGAGGCAGATTCA	1149
Db	361	TCTGAGGTGGAAATTTATTCATGTTTACATCCCTTATCTGTTGAAAACAAGGAGGCAGATTCA	420
QY	1150	TTCCGTTATCTCTAAAAACAGGTACAGCAAAATCTCTAAAGTCACTTTTAAAGTGTCAAGAAATA	1209
Db	421	TTCCGTTATCTCTAAAAACAGGTACAGCAAAATCTCTAAAGTCACTTTTAAAGTGTCAAGAAATA	480
QY	1210	ATGATTGATGCTGAAGGAAGATCATAGATGTCATAGATAAGCAACTAATCTCAACTCTTT	1269
Db	481	ATGATTGATGCTGAAGGAAGATCATAGATGTCATAGATAAGCAACTAATCTCAACTCTTT	540
QY	1270	GAGATTTCTATTGGAAGGATTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGAAAA	1329
Db	541	GAGATTTCTATTGGAAGGATTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGAAAA	600
QY	1330	TATGCTTGGTCCATCTCTACTAGATCGCTCCCA	1361
Db	601	TATGCTTGGTCCATCTCTACTAGATCGCTCCCA	632
RESULT 49			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
QY	730	GTAAAGATGAGGGCCCAAGGATTTAGCAACACCTTTAAGGCCCAATCTAGTGGAA	789
Db	1	GTAAAGATGAGGGCCCAAGGATTTAGCAACACCTTTAAGGCCCAATCTAGTGGAA	60
QY	790	ACTAGTTGTCTCCAAACATACCGATGGATCCAAATTTATGCGCGCTGATCCAGATGGATT	849
Db	61	ACTAGTTGTCTCCAAACATACCGATGGATCCAAATTTATGCGCGCTGATCCAGATGGATT	120
QY	850	GCTTTTATACATCAACGATATTTGGATATCTAACATCTAACACGAGAGAAAGGAGA	909
Db	121	GCTTTTATACATCAACGATATTTGGATATCTAACATCTAACACGAGAGAAAGGAGA	180
QY	910	CTCATTTATGTGCAATAGTAGTCCCAACATGGAAGAGATCCAGATTCAGCTGGAGTC	969
Db	181	CTCATTTATGTGCAATAGTAGTCCCAACATGGAAGAGATCCAGATTCAGCTGGAGTC	240
QY	970	GCTACCTTTGTCTCCAAAGAGATTTGATAGATTTCTGGCTATTTGGTGGTGTCCAAA	1029
Db	241	GCTACCTTTGTCTCCAAAGAGATTTGATAGATTTCTGGCTATTTGGTGGTGTCCAAA	300
QY	1030	GCTGAACAACTCCCACTGGTGGTAAAAATTTCTAGAAATTTCTATATGAAGAAATGATGAA	1089
Db	301	GCTGAACAACTCCCACTGGTGGTAAAAATTTCTAGAAATTTCTATATGAAGAAATGATGAA	360
QY	1090	TCTGAGGTGGAAATTTATTCATGTTTACATCCCTTATCTGTTGAAAACAAGGAGGCAGATTCA	1149
Db	361	TCTGAGGTGGAAATTTATTCATGTTTACATCCCTTATCTGTTGAAAACAAGGAGGCAGATTCA	420
QY	1150	TTCCGTTATCTCTAAAAACAGGTACAGCAAAATCTCTAAAGTCACTTTTAAAGTGTCAAGAAATA	1209
Db	421	TTCCGTTATCTCTAAAAACAGGTACAGCAAAATCTCTAAAGTCACTTTTAAAGTGTCAAGAAATA	480
QY	1210	ATGATTGATGCTGAAGGAAGATCATAGATGTCATAGATAAGCAACTAATCTCAACTCTTT	1269
Db	481	ATGATTGATGCTGAAGGAAGATCATAGATGTCATAGATAAGCAACTAATCTCAACTCTTT	540
QY	1270	GAGATTTCTATTGGAAGGATTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGAAAA	1329
Db	541	GAGATTTCTATTGGAAGGATTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGAAAA	600
QY	1330	TATGCTTGGTCCATCTCTACTAGATCGCTCCCA	1361
Db	601	TATGCTTGGTCCATCTCTACTAGATCGCTCCCA	632
RESULT 49			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
QY	730	GTAAAGATGAGGGCCCAAGGATTTAGCAACACCTTTAAGGCCCAATCTAGTGGAA	789
Db	1	GTAAAGATGAGGGCCCAAGGATTTAGCAACACCTTTAAGGCCCAATCTAGTGGAA	60
QY	790	ACTAGTTGTCTCCAAACATACCGATGGATCCAAATTTATGCGCGCTGATCCAGATGGATT	849
Db	61	ACTAGTTGTCTCCAAACATACCGATGGATCCAAATTTATGCGCGCTGATCCAGATGGATT	120
QY	850	GCTTTTATACATCAACGATATTTGGATATCTAACATCTAACACGAGAGAAAGGAGA	909
Db	121	GCTTTTATACATCAACGATATTTGGATATCTAACATCTAACACGAGAGAAAGGAGA	180
QY	910	CTCATTTATGTGCAATAGTAGTCCCAACATGGAAGAGATCCAGATTCAGCTGGAGTC	969
Db	181	CTCATTTATGTGCAATAGTAGTCCCAACATGGAAGAGATCCAGATTCAGCTGGAGTC	240
QY	970	GCTACCTTTGTCTCCAAAGAGATTTGATAGATTTCTGGCTATTTGGTGGTGTCCAAA	1029
Db	241	GCTACCTTTGTCTCCAAAGAGATTTGATAGATTTCTGGCTATTTGGTGGTGTCCAAA	300
QY	1030	GCTGAACAACTCCCACTGGTGGTAAAAATTTCTAGAAATTTCTATATGAAGAAATGATGAA	1089
Db	301	GCTGAACAACTCCCACTGGTGGTAAAAATTTCTAGAAATTTCTATATGAAGAAATGATGAA	360

	Matches	580;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	5	GCTAAGCCTCCGAGGCCAAGGCGCTGCTACTACGCCGCGCTGCTTCTTAGTGCCGCGTT	64							
Db	1	GCTAAGCCTCCGAGGCCAAGGCGCTGCTACTACGCCGCGCTGCTTCTTAGTGCCGCGTT	60							
Qy	65	CGCGCCTGGGTTGTTCACCGGCGCGCGCGCGAGGAAGCACTGCACACAGAGACCGAGT	124							
Db	61	CGCGCCTGGGTTGTTCACCGGCGCGCGCGAGGAAGCACTGCACACAGAGACCGAGT	120							
Qy	125	GGAGCGCGCGCACATGACGAGCGCGCGCGCTCCATAGCGCACGTCGGAGCGGTCG	184							
Db	121	GGAGCGCGCGCACATGACGAGCGCGCGCGCTCCATAGCGCACGTCGGAGCGGTCG	180							
Qy	185	GGCGGGCGGGGGAGGAAATGCAACATGCGACGACGCAATGGAAACAGAAACAGCTGG	244							
Db	181	GGCGGGCGGGGGAGGAAATGCAACATGCGACGACGCAATGGAAACAGAAACAGCTGG	240							
Qy	245	GTGTTGAGATATTGAAACTCGGACCTGTGAGGAGAAATATGAAATCAGAGATCGGCCTA	304							
Db	241	GTGTTGAGATATTGAAACTCGGACCTGTGAGGAGAAATATGAAATCAGAGATCGGCCTA	300							
Qy	305	AATTGGAGCCTTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGACTGCTGCCG	364							
Db	301	AATTGGAGCCTTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGACTGCTGCCG	360							
Qy	365	ATACCAAGAAATATCATGGCTACATGATGCTTAAGCCACCATGATTTCACTTTGTGA	424							
Db	361	ATACCAAGAAATATCATGGCTACATGATGCTTAAGCCACCATGATTTCACTTTGTGA	420							
Qy	425	AGAGAAATGATCCAGATGGACCTCATTCACAGACAGATCTATTACTTGGCATGCTGGTG	484							
Db	421	AGAGAAATGATCCAGATGGACCTCATTCACAGACAGATCTATTACTTGGCATGCTGGTG	480							
Qy	485	AGAACAGAGAAATACACTGTTTATTTCTGAAATGCCAAACTATCAATAGACGACGAG	544							
Db	481	AGAACAGAGAAATACACTGTTTATTTCTGAAATGCCAAACTATCAATAGACGACGAG	540							
Qy	545	TCCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTCA	584							
Db	541	TCCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTCA	580							

```

RESULT 50
BI084090
LOCUS
DEFINITION
        BI084090 1041 bp mRNA linear EST 20-JUN-2001
        602869453F1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5013996 5',
        mRNA sequence.
ACCESSION
        BI084090
VERSION
        BI084090.1 GI:14502420
KEYWORDS
        EST.
SOURCE
        Homo sapiens (human)
        Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
        Homnidae; Homo
REFERENCE
        1 (bases 1 to 1041)
        NIH-MGC http://mgc.nci.nih.gov/.
        National Institutes of Health, Mammalian Gene Collection (MGC)
        Unpublished (1999)
        Contact: Robert Straubeberg, Ph.D.
        Email: cgapbs-re@mail.nih.gov
        Tissue Procurement: ATCC
        CDNA Library Preparation: CLONETECH Laboratories, Inc.
        CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
        DNA Sequencing by: Incyte Genomics, Inc.
        Clone distribution: MGC clone distribution information can be
        found through the I.M.A.G.E. Consortium/LNL at:
        http://image.lnl.gov
        Plate: L12CM1820 row: j column: 13
        High quality sequence stop: 847.
        Location/Qualifiers
            1..1041
FEATURES
        source

```

FEATURES
SOURCE

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:501396"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/clone_lib="NIH_MGC_102"
/notes="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match      18.6%; Score 580; DB 2; Length 1041;
Best Local Similarity 100.0%; Pred. No. 6.7e-302;
Matches 580; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1631 GCAATATATAACGATCCAGTGGTGGGCTGCTCTCCAAAGTATTTCAAGTGTCTCATCA 1690
DB 1 GCAATATATAACGATCCAGTGGTGGGCTGCTCTCCAAAGTATTTCAAGTGTCTCATCA 60

QY 1691 AAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATA 1750
DB 61 AAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATA 120

QY 1751 TCCAAGTTGATGAAGTCAGAGGCTGGTATATTTGAAGGCCACCAAGAGACTCCCTTTAG 1810
DB 121 TCCAAGTTGATGAAGTCAGAGGCTGGTATATTTGAAGGCCACCAAGAGACTCCCTTTAG 180

QY 1811 AGCATCACCTGTACGTAGTCAGTTAGTAAATCTCGGAGAGGTGACAGGCTGACTGACC 1870
DB 181 AGCATCACCTGTACGTAGTCAGTTAGTAAATCTCGGAGAGGTGACAGGCTGACTGACC 240

QY 1871 GTGGCTACTCAATCTTCTGCTGCATGATGACGACACTGTGACTCTCTTATAAGTAAGTATA 1930
DB 241 GTGGCTACTCAATCTTCTGCTGCATGATGACGACACTGTGACTCTCTTATAAGTAAGTATA 300

QY 1931 GTAACCCAGAAATCCACTGTGTGTCCTTTTACAAGCTATCAAGTCTTGAAGATGACC 1990
DB 301 GTAACCCAGAAATCCACTGTGTGTCCTTTTACAAGCTATCAAGTCTTGAAGATGACC 360

QY 1991 CAATCTTGAAAACAAAGGAATTTTGGGCCACATTTTGGGATTCAGCAGGCTCTCTTCCTG 2050
DB 361 CAATCTTGAAAACAAAGGAATTTTGGGCCACATTTTGGGATTCAGCAGGCTCTCTTCCTG 420

QY 2051 ACTATATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGATTTACATGTATGGGA 2110
DB 421 ACTATATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGATTTACATGTATGGGA 480

QY 2111 TGCTCTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTCATAT 2170
DB 481 TGCTCTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTCATAT 540

QY 2171 ATGTGGTGCCTCAGGTGAGTGTGTAATATCGGTTTAA 2210
DB 541 ATGTGGTGCCTCAGGTGAGTGTGTAATATCGGTTTAA 580

RESULT 51
DB156844
LOCUS
DEFINITION
DB156844
THYMU3 Homo sapiens cDNA clone THYMU3035709 5', mRNA
sequence.
ACCESSION
DB156844
VERSION
DB156844.1 GI:83187098
KEYWORDS
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 593)
 Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
 Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
 Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
 Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
 Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
 Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
 Diversification and Characterization of Putative Alternative
 Promoters of Human Genes
 Genome Res. 16 (1), 55-65 (2006)
 16344560
 Contact: Takao Isogai
 FLJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: flj-cdna@nifty.com
 NEDO human cDNA project (New Energy and Industrial Technology
 Developmental Organization, Japan); cDNA library construction;
 Helix Research Institute (HRI); 5'-end one pass sequencing; HRI,
 Research Association for Biotechnology (RAB) and Biotechnology
 Center, National Institute of Technology and Evaluation; 3'-end one
 pass sequencing; RAB.
 Location/Qualifiers
 1..593
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="THYMU3035709"
 /tissue_type="thymus"
 /clone_lib="THYMU3"
 /note="Vector: pME18SFL3"

ORIGIN
 Query Match 18.5%; Score 578; DB 9; Length 593;
 Best Local Similarity 100.0%; Pred. No. 7.9e-301;
 Matches 578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAGCTCCGAGCGCAAGCGCTGCTACTGCGCGCGCTGCTTCTAGTCCGCGTCCG 67
 DB 2 AAGCTCCGAGCGCAAGCGCTGCTACTGCGCGCGCTGCTTCTAGTCCGCGTCCG 61
 QY 68 CGCTTGGGTGTGTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGTGGG 127
 DB 62 CGCTTGGGTGTGTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGTGGG 121
 QY 128 GCGCGCGCAGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 187
 DB 122 GCGCGCGCAGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 181
 QY 188 GCG 247
 DB 182 GCG 241
 QY 248 TTGAGATATTCGAACTGGGACTGTGAGGAGATATTCGATCAGGATCGCGCTAAT 307
 DB 242 TTGAGATATTCGAACTGGGACTGTGAGGAGATATTCGATCAGGATCGCGCTAAT 301
 QY 308 TGAGCGCTTTTATTTGTTAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATA 367
 DB 302 TGAGCGCTTTTATTTGTTAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATA 361
 QY 368 CCGAATAATCATGCTCATGATGGCTTAAGGACCAATGATTTTCATGTTGTGAGG 427
 DB 362 CCGAATAATCATGCTCATGATGGCTTAAGGACCAATGATTTTCATGTTGTGAGG 421
 QY 428 GGAATGATCCAGATGACCTCAATTCAGACAGAACTTATTCCTGCGCATGCTGCGGTGAGA 487
 DB 422 GGAATGATCCAGATGACCTCAATTCAGACAGAACTTATTCCTGCGCATGCTGCGGTGAGA 481

QY 488 ACAGAGAAAATACACTGTTTATTTCTGAAATCCCAAACTATCAATAGAGCAGAGTCT 547
 DB 482 ACAGAGAAAATACACTGTTTATTTCTGAAATCCCAAACTATCAATAGAGCAGAGTCT 541
 QY 548 TAATGCTCTCTTGGAGAGCCTCTTTTGGATCTTTTTCAG 585
 DB 542 TAATGCTCTCTTGGAGAGCCTCTTTTGGATCTTTTTCAG 579

RESULT 52
 LOCUS DA140770 629 bp mRNA linear EST 31-OCT-2005
 DEFINITION DA140770 BRALZ2 Homo sapiens cDNA clone BRALZ2015423 5', mRNA
 sequence.
 ACCESSION DA140770
 VERSION DA140770
 KEYWORDS DA140770.1 GI:78458395
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 629)
 Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
 Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
 Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
 Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
 Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
 Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
 Diversification and Characterization of Putative Alternative
 Promoters of Human Genes
 Genome Res. 16 (1), 55-65 (2006)
 16344560
 Contact: Takao Isogai
 FLJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: flj-cdna@nifty.com
 NEDO human cDNA project (New Energy and Industrial Technology
 Developmental Organization, Japan); cDNA library construction;
 Helix Research Institute (HRI); 5'-end one pass sequencing; HRI,
 Research Association for Biotechnology (RAB) and Biotechnology
 Center, National Institute of Technology and Evaluation; 3'-end one
 pass sequencing; RAB.
 Location/Qualifiers
 1..629
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="BRALZ2015423"
 /tissue_type="alzheimer cortex"
 /clone_lib="BRALZ2"
 /note="Vector: pME18SFL3"

ORIGIN
 Query Match 18.5%; Score 578; DB 9; Length 629;
 Best Local Similarity 99.8%; Pred. No. 7.9e-301;
 Matches 628; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 462 CTATTACCTTGGCCATGTCTGGTGAGAAACAGAGAAATACACTGTTTATTCGAAATCC 521
 DB 1 CTATTACCTTGGCCATGTCTGGTGAGAAACAGAGAAATACACTGTTTATTCGAAATCC 60
 QY 522 CAAAACTATCAATAGAGCAGCTTAAATGCTCTTGGAGCCTCTTTGGATCTTTT 581
 DB 61 CAAAACTATCAATAGAGCAGCTTAAATGCTCTTGGAGCCTCTTTGGATCTTTT 120
 QY 582 TCAGGCAACACTGAGCTATGGAATGATTTCTCGAGAGAGAACTATTAAGAGAGAAA 641

```
Db 121 TCAGGCAACACTGGCACTATGGAATGATTTCTCGAGAGAGAACTATTAAAGAGAAAGAAA 180
Qy 642 ACGCATTTGGAACTAGTCGGAATTCCTTTACCATATTACCAAGAGAGTGAACATTTCT 701
Db 181 ACGCATTTGGAACTAGTCGGAATTCCTTTACCATATTACCAAGAGAGTGAACATTTCT 240
Qy 702 GTTTCAGCCGGTAGTGGAAATTTATCAGTAAAGATGGAGGCGCACAGATTTACGCA 761
Db 241 GTTTCAGCCGGTAGTGGAAATTTATCAGTAAAGATGGAGGCGCACAGATTTACGCA 300
Qy 762 ACAACCTTTAAGGCCCAATCTAGTGAAGAACTAGTTGTCCCAACATACGGATGATCCAAA 821
Db 301 ACAACCTTTAAGGCCCAATCTAGTGAAGAACTAGTTGTCCCAACATACGGATGATCCAAA 360
Qy 822 ATTATGCCCGGTGATCCAGACTGGATTGCTTTATACATAGCAACATATTGGATATC 881
Db 361 ATTATGCCCGGTGATCCAGACTGGATTGCTTTATACATAGCAACATATTGGATATC 420
Qy 882 TAACATCGTAACACAGAGAGAAAGAGACTCACTATTGTGCAACATGAGCTAGCCAAAT 941
Db 421 TAACATCGTAACACAGAGAGAAAGAGACTCACTATTGTGCAACATGAGCTAGCCAAAT 480
Qy 942 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTTCCCAAGAGAAATTTGATAG 1001
Db 481 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTTCCCAAGAGAAATTTGATAG 540
Qy 1002 ATATTCTGGCTATTGGTGTGTCMAAAGCTGAACAACTCCAGTGGTGTGTAATAATCT 1061
Db 541 ATATTCTGGCTATTGGTGTGTCMAAAGCTGAACAACTCCAGTGGTGTGTAATAATCT 600
Qy 1062 TAGAATTCATATGAAGAAATGATGAAT 1090
Db 601 TAGAATTCATATGAAGAAATGATGAAT 629

RESULT 53
DA301389
LOCUS
DEFINITION DA301389 BRHIP2 Homo sapiens cDNA clone BRHIP201610 5', mRNA linear EST 03-NOV-2005
sequence.
ACCESSION DA301389
VERSION DA301389.1 GI:79169476
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo
REFERENCE 1 (bases 1 to 576)
Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
Yamashita,P., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
Tanase,I., Negai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
1634560
Contract: Takao Isogai
FLN Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdn@nifty.com
NEO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction;
Helix Research Institute (HRI); 5'-end one pass sequencing; HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
```

FEATURES

source

Location/Qualifiers

1..576
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRHIP201610"
/tissue_type="hippocampus"
/clone_lib="BRHIP2"
/note="Vector: pMB18SFL3"

ORIGIN

Query Match 18.5%; Score 576; DB 9; Length 576;
Best Local Similarity 100.0%; Pred. No. 9.6e-300; Indels 0; Gaps 0;
Matches 576; Conservative 0; Mismatches 0;

```
Qy 14 TCAGAGGCGCAAGGCGCGTCTACTGCGCGCGCTGCTTCTTAGTGCGCGCTTCCGCGCTG 73
Db 1 TCAGAGGCGCAAGGCGCGTCTACTGCGCGCGCTGCTTCTTAGTGCGCGCTTCCGCGCTG 60
Qy 74 GGTTCATACCAGCGCGCGCGCGCGAGGAGCACTGCAACACAGGACCGGAGTGGAGCGCG 133
Db 61 GGTTCATACCAGCGCGCGCGCGCGAGGAGCACTGCAACACAGGACCGGAGTGGAGCGCG 120
Qy 134 GCAGCATGAAGCGCGCGCGCGCGCTCCATAGCGGACGTCGGGAGCGTCCGCGCGCGCGCC 193
Db 121 GCAGCATGAAGCGCGCGCGCGCGCTCCATAGCGGACGTCGGGAGCGTCCGCGCGCGCGCC 180
Qy 194 GGGGGGAGGAAATGCAACATGGCAGCAGCAATGGAACACAGACAGCTGGTGTGAGA 253
Db 181 GGGGGGAGGAAATGCAACATGGCAGCAGCAATGGAACACAGACAGCTGGTGTGAGA 240
Qy 254 TATTGAACTGCGGACTGTGAGGAGAAATATTGAATCAGAGATCGGCTTAATTTGGAGC 313
Db 241 TATTGAACTGCGGACTGTGAGGAGAAATATTGAATCAGAGATCGGCTTAATTTGGAGC 300
Qy 314 CTTTATGTTGAGCGGTATTCCTGGAGTTCAGCTTAAAGCTCTTTCGCGATACAGAA 373
Db 301 CTTTATGTTGAGCGGTATTCCTGGAGTTCAGCTTAAAGCTCTTTCGCGATACAGAA 360
Qy 374 AATATCATGGCTACATGATGCTAAGGACCAACATGATTTCAATGTTGTGAAGGAATG 433
Db 361 AATATCATGGCTACATGATGCTAAGGACCAACATGATTTCAATGTTGTGAAGGAATG 420
Qy 434 ATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGCATGCTGTGTGAGAACAG 493
Db 421 ATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGCATGCTGTGTGAGAACAG 480
Qy 494 AAAATACACTGTTTATTCTGAATTCCTCAAACTATCAATAGCAGCAGCTTTAATGC 553
Db 481 AAAATACACTGTTTATTCTGAATTCCTCAAACTATCAATAGCAGCAGCTTTAATGC 540
Qy 554 TCTCTTGAAGCGCTCTTTTGGATCTTTTCAGGCAA 589
Db 541 TCTCTTGAAGCGCTCTTTTGGATCTTTTCAGGCAA 576
```

RESULT 54

DA690898

LOCUS

DEFINITION DA690898 NT2NE2 Homo sapiens cDNA clone NT2NE2010446 5', mRNA linear EST 07-NOV-2005 sequence.

ACCESSION DA690898

VERSION DA690898.1

KEYWORDS GI:81154735

SOURCE EST.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo

1 (bases 1 to 576)

Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,

Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,

Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,

REFERENCE

AUTHORS

Thu Jun 22 09:04:26 2006

Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S. Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes

Genome Res. 16 (1), 55-65 (2006)

16344560

CONTACT: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

Location/Qualifiers

1. .576
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2NE2010446"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_lib="NT2NE2"
/notes="Vector: pME18SFL3; mRNA from NT2 neuron after the differentiation of NT2 neuronal precursor cells"

ORIGIN

Query Match 18.5%; Score 576; DB 9; Length 576;
Best Local Similarity 100.0%; Pred. No. 9.6e-300;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

222 AGCAATGGAACAGAACAGCTGGTCTTGGATATTTGAACCTGGGACTGTGAGGAGAA 281
|||||
1 AGCAATGGAACAGAACAGCTGGTCTTGGATATTTGAACCTGGGACTGTGAGGAGAA 60
|||||

282 TATTGAATTCACAGATCGGCTAAATTTGGAGCGCTTTTATGTTGAGCGGTATTCCTGGAG 341
|||||

61 TATTGAATTCACAGATCGGCTAAATTTGGAGCGCTTTTATGTTGAGCGGTATTCCTGGAG 120
|||||

342 TCAGCTTAAAGAGCTGCTGGCATACAGAAATATCATGCTACATGCTGCTTAAGGC 401
|||||

121 TCAGCTTAAAGAGCTGCTGGCATACAGAAATATCATGCTACATGCTGCTTAAGGC 180
|||||

402 ACCACATGATTTTCATGTTTGTGAAGAGGATGATCCAGATGGACCTCATTCAGACAGAA 461
|||||

181 ACCACATGATTTTCATGTTTGTGAAGAGGATGATCCAGATGGACCTCATTCAGACAGAA 240
|||||

462 CTATTAACCTTGCCATGCTGCTGGAGAACAGAGAAATACACTGTTTATTCGAAATCC 521
|||||

241 CTATTAACCTTGCCATGCTGCTGGAGAACAGAGAAATACACTGTTTATTCGAAATCC 300
|||||

522 CAAAACTATCAATAGACAGCAGTCTTAATGCTCTCTTGGAGCGCTTTTGGATCTTTT 581
|||||

301 CAATACTATCAATAGACAGCAGTCTTAATGCTCTCTTGGAGCGCTTTTGGATCTTTT 360
|||||

582 TCAGGCAACACTGACATGATGATGATTTCTCGAAGAGAACTATTAAGAGAGAGAA 641
|||||

361 TCAGGCAACACTGACATGATGATGATTTCTCGAAGAGAACTATTAAGAGAGAGAA 420
|||||

642 ACGATTGGAACAGCTCGGAATTCGTTCTTACGATATTCACAGAGAGTGAACATTCT 701
|||||

421 ACGATTGGAACAGCTCGGAATTCGTTCTTACGATATTCACAGAGAGTGAACATTCT 480
|||||

702 GTTTCACAGCGGTAGTGGAAATTTATTCAGTAAAGATGGAGGCCACAGAGATTACGCA 761
|||||

481 GTTTCACAGCGGTAGTGGAAATTTATTCAGTAAAGATGGAGGCCACAGAGATTACGCA 540
|||||

762 ACAACCTTTAAGCCCAATCTAGTGGAACTAGTTG 797
|||||

541 ACAACCTTTAAGCCCAATCTAGTGGAACTAGTTG 576
|||||

RESULT 55
DA478474 577 bp mRNA linear EST 06-NOV-2005
LOCUS DA478474 DFNES2 Homo sapiens cDNA clone DFNES2008211 5', mRNA
DEFINITION DA478474 sequence.
ACCESSION DA478474
VERSION DA478474.1 GI:80934675
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 577)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S. Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
CONTACT: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

Location/Qualifiers

1. .577
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DFNES2008211"
/cell_type="normal dermal fibroblasts (Neonatal Skin) (NHDP2564)"
/clone_lib="DFNES2"
/note="Vector: pME18SFL3; primary culture, normal dermal fibroblasts"

FEATURES
source

ORIGIN

Query Match 18.5%; Score 576; DB 9; Length 577;
Best Local Similarity 100.0%; Pred. No. 9.6e-300; Indels 0; Gaps 0;
Matches 576; Conservative 0; Mismatches 0;

1 AAGTGCTAAGCTCCGAGGCCAAGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCCG 60
|||||

2 AAGTGCTAAGCTCCGAGGCCAAGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCCG 61
|||||

61 CGTTTCGCGCGCTCGGTTGTCAACCGCGCGCGCGCGAGGAGCACTGCAACACGAGCCG 120
|||||

62 CGTTTCGCGCGCTCGGTTGTCAACCGCGCGCGCGCGAGGAGCACTGCAACACGAGCCG 121
|||||

121 GAGTGAGAGCGCGCGAGCATGAGCGCGCGCGCGCTTCCATAGCGCACTGCGGAGCG 180
|||||

122 GAGTGAGAGCGCGCGAGCATGAGCGCGCGCGCGCTTCCATAGCGCACTGCGGAGCG 181
|||||

```

181 TCCGGGCGGGCCGGGGGAGGAAATGCAATGCGCAGCAGCAATGCAACAGACAG 240
182 TCCGGGCGGGCCGGGGGAGGAAATGCAACATGCGCAGCAGCAATGCAACAGACAG 241
241 CTGGGCTTGGAGATATTTGAACTCGGAGCTGTGAGGAGATATTTGAATCAGAGATCGG 300
242 CTGGGCTTGGAGATATTTGAACTCGGAGCTGTGAGGAGATATTTGAATCAGAGATCGG 301
301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTTCTGGAGCTGAGCTTAAAAAGCTGCTT 360
302 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTTCTGGAGCTGAGCTTAAAAAGCTGCTT 361
361 GCCGATACCCAGAAATATCATGCTACATGATGCTTAAGGCAACACATGATTTTCAATGTTT 420
362 GCCGATACCCAGAAATATCATGCTACATGATGCTTAAGGCAACACATGATTTTCAATGTTT 421
421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGATCTATTACCTTCCCATGCT 480
422 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGATCTATTACCTTCCCATGCT 481
481 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGATCTATTACCTTCCCATGCT 480
482 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGATCTATTACCTTCCCATGCT 540
541 GCAGTCTTAATGCTCTCTTGAAGCTCTTTGGAT 576
542 GCAGTCTTAATGCTCTCTTGAAGCTCTTTGGAT 577

```

```

RESULT 56
BG752328
LOCUS
DEFINITION
602730802P1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4874424 5',
mRNA sequence.
BG752328
BG752328.1 GI:14062981
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo
1 (bases 1 to 627)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L10M1754 row: c column: 01
High quality sequence stop: 627.
Location/Qualifiers
1..627
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4874424"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="PH10B (phage-resistant)"
/clone_lib="NIH MGC 43"
/note="Organ: eye; Vector: pOTB7; Site: 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACAGAG(C). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

```

```

181 TCCGGGCGGGCCGGGGGAGGAAATGCAATGCGCAGCAGCAATGCAACAGACAG 240
182 TCCGGGCGGGCCGGGGGAGGAAATGCAACATGCGCAGCAGCAATGCAACAGACAG 241
241 CTGGGCTTGGAGATATTTGAACTCGGAGCTGTGAGGAGATATTTGAATCAGAGATCGG 300
242 CTGGGCTTGGAGATATTTGAACTCGGAGCTGTGAGGAGATATTTGAATCAGAGATCGG 301
301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTTCTGGAGCTGAGCTTAAAAAGCTGCTT 360
302 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTTCTGGAGCTGAGCTTAAAAAGCTGCTT 361
361 GCCGATACCCAGAAATATCATGCTACATGATGCTTAAGGCAACACATGATTTTCAATGTTT 420
362 GCCGATACCCAGAAATATCATGCTACATGATGCTTAAGGCAACACATGATTTTCAATGTTT 421
421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGATCTATTACCTTCCCATGCT 480
422 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGATCTATTACCTTCCCATGCT 481
481 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGATCTATTACCTTCCCATGCT 480
482 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGATCTATTACCTTCCCATGCT 540
541 GCAGTCTTAATGCTCTCTTGAAGCTCTTTGGAT 576
542 GCAGTCTTAATGCTCTCTTGAAGCTCTTTGGAT 577

```

Note: this is a NIH_MGC Library. |"

```

Query Match 18.5%; Score 576; DB 2; Length 627;
Best Local Similarity 99.8%; Pred. No. 9.6e-300;
Matches 626; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2412 GGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGCATTAATGAGGTC 2471
Db 1 GGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGCATTAATGAGGTC 60
QY 2472 AGATATCTTCACGGGTTGCTATTGCTGGGGCCCCCAGTCACCTCTGATGATAC 2531
Db 61 AGATATCTTCACGGGTTGCTATTGCTGGGGCCCCCAGTCACCTCTGATGATAC 120
QY 2532 AGATATCTTCACGGGTTGCTATTGCTGGGGCCCCCAGTCACCTCTGATGATAC 2591
Db 121 AGATATCTTCACGGGTTGCTATTGCTGGGGCCCCCAGTCACCTCTGATGATAC 180
QY 2592 ATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAATCGTTTACTGCTTTACA 2651
Db 181 ATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAATCGTTTACTGCTTTACA 240
QY 2652 TGGTTTCTGATGAGAAATGTCATTTTGCATACCACTATTTACTGATTTTACT 2711
Db 241 TGGTTTCTGATGAGAAATGTCATTTTGCATACCACTATTTACTGATTTTACT 300
QY 2712 GAGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGAGAGACACAGATAAGT 2771
Db 301 GAGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGAGAGACACAGATAAGT 360
QY 2772 TCCTGAATCGGAGAACATTATGAACTGCTTTTGCATCTTCAAGAAAAACCTTGG 2831
Db 361 TCCTGAATCGGAGAACATTATGAACTGCTTTTGCATCTTCAAGAAAAACCTTGG 420
QY 2832 ATCACTGTTTCTGCTCTTAAAGTGATATTTTGAACCTGTGTGAGAACTCTCTG 2891
Db 421 ATCACTGTTTCTGCTCTTAAAGTGATATTTTGAACCTGTGTGAGAACTCTCTG 480
QY 2892 CACTGGCTATTTAAACCAATGAGGAGTTTATCAACAGAAAAACAGAAATTCATCA 2951
Db 481 CACTGGCTATTTAAACCAATGAGGAGTTTATCAACAGAAAAACAGAAATTCATCA 540
QY 2952 CATTTTGATACCTGCCATGATCACTCTCTGAAAAATAATGTTGCTGAGGGG 3011
Db 541 CATTTTGATACCTGCCATGATCACTCTCTGAAAAATAATGTTGCTGAGGGG 600
QY 3012 TCTACGGTTTCTGCTGATGATCTAATA 3038
Db 601 TCTACGGTTTCTGCTGATGATCTAATA 627

```

```

RESULT 57
DAL63570
LOCUS
DEFINITION
DAL63570 BRAWY2 Homo sapiens cDNA clone BRAMY2024215 5', mRNA
sequence.
ACCESSION
DAL63570
VERSION
DAL63570.1 GI:78268098
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo
1 (bases 1 to 575)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Teuritani, K., Wakaquri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kuchida, N.,
Yonekura, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Watanabe, M.,
Muraoka, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Negai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale

```


QY	AAAGCTGCTTGGCGATACACAGAAAAATATCATGGCTACATGATGGCTTAAGCCACCATGA	410
Db	AAAGCTGCTTGGCGATACACAGAAAAATATCATGGCTACATGATGGCTTAAGCCACCATGA	360
QY	TTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTTATTACCT	470
Db	TTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTTATTACCT	420
QY	TGCCATGCTCTGGTGAAGACAGAGAAATACACTGTGTTTATCTCGAAATCCCAAACTAT	530
Db	TGCCATGCTCTGGTGAAGACAGAGAAATACACTGTGTTTATCTCGAAATCCCAAACTAT	480
QY	CAATAGACAGACAGCTCTTATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTCAGGGCAAC	590
Db	CAATAGACAGACAGCTCTTATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTCAGGGCAAC	540
QY	ACTGACTATGGAATGATTTCTCGAAGAAGA	623
Db	ACTGACTATGGAATGATTTCTCGAAGAAGA	573
RESULT 59		
DA931907		
LOCUS		
DEFINITION	DA931907 SPLEN2 Homo sapiens cDNA clone SPLEN2001115 5', mRNA	linear EST 11-NOV-2005
ACCESSION	DA931907	sequence
VERSION	DA931907.1	GI:82063318
KEYWORDS	EST	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo	
AUTHORS	Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Teuricani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saito,K., Iseno,Y., Irie,R., Kushida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Megatsuma,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.	
TITLE	Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes	
JOURNAL	Genome Res. 16 (1), 55-65 (2006)	
PUBMED	16344560	
COMMENT	Contact: Takao Isogai FLJ Project (HRI Team) Helix Research Institute 2-6-7 Kausa-Kamatari, Kisarazu, Chiba, 292-0818, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: flj-cdna@nifty.com NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); S'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB	
FEATURES	Location/Qualifiers	
source	1..574	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="SPLEN2001115"	
	/tissue_type="spleen"	
	/clone_lib="SPLEN2"	
	/note="Vector: pWE18SPL3"	
ORIGIN		
Query Match	18.4%; Score 573; DB 9; Length 574;	
Best Local Similarity	100.0%; Pred. No. 4.1e-298;	

us-10-825-632-2.olig.rst

Thu Jun 22 09:04:26 2006

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 760)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1041 Std Error: 0.00
Seq Primer: -40UP from Gibco
High quality sequence stop: 444.

FEATURES
Location/Qualifiers
1..760
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2240461"
/issue_type="pooled germ cell tumors"
/lab_host="PH10B"
/clone_lib="NCI CGAP GC6"
/notes="Vector: pTT3D-PacI; Site 1: Not 1; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP GC4 was
prepared, and as circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(clones 1257096-1258631, 1469064-1470983, and
1475592-1476743). Subtraction by Bento Soares and M.
Fatima Bonaldo."

ORIGIN
Query Match 18.0%; Score 562; DB 1; Length 760;
Best Local Similarity 99.8%; Pred. No. 3.8e-292;
Matches 612; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2503 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAGCTTATATGGGTCAACCT 2562
DB 613 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAGCTTATATGGGTCAACCT 554
QY 2563 GACCAAGATGACAGGGCTTATCTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCC 2622
DB 553 GACCAAGATGACAGGGCTTATCTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCC 494
QY 2623 TCTGAACCAAAATCGTTTACTCTTATCATGCTTCTTCTGGATGAGATGTCCTATTTTGA 2682
DB 493 TCTGAACCAAAATCGTTTACTCTTATCATGCTTCTTCTGGATGAGATGTCCTATTTGCA 434
QY 2683 CATACAGATATATCTAGATGTTTATGAGGGCTGGAAAGCCATATGATGATACAGATC 2742
DB 433 CATACAGATATATCTAGATGTTTATGAGGGCTGGAAAGCCATATGATGATACAGATC 374
QY 2743 TATCCTCAGGAGAGACACAGATTAAGATTTCTGATCGGAGAACATATGAACTGCAT 2802
DB 373 TATCCTCAGGAGAGACACAGATTAAGATTTCTGATCGGAGAACATATGAACTGCAT 314
QY 2803 CTTTGGACCTACCTTCAAGAAACCTTGGATCACGTTATTTGCTCTTAAAGTATATAA 2862
DB 313 CTTTGGACCTACCTTCAAGAAACCTTGGATCACGTTATTTGCTCTTAAAGTATATAA 254
QY 2863 TTTTGGACCTGTGTAAGAACTCTCTGGATATACATGGCTTATTAACCAATAGAGAGTTTA 2922

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 760)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1041 Std Error: 0.00
Seq Primer: -40UP from Gibco
High quality sequence stop: 444.

FEATURES
Location/Qualifiers
1..760
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE7010347"
/issue_type="placenta"
/clone_lib="PLACE7"
/note="Vector: pME18SFL3"

ORIGIN
Query Match 18.3%; Score 571; DB 9; Length 573;
Best Local Similarity 100.0%; Pred. No. 4.9e-297; Indels 0; Gaps 0;
Matches 571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGTCTAAGCTCGAGCCCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGGCG 60
DB 3 AAGTCTAAGCTCGAGCCCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGGCG 62
QY 61 CGTTTCGCGCGCTGGTGTGTACCGGCGCGCGCGCGCGAGCCACCAACGAGCGG 120
DB 63 CGTTTCGCGCGCTGGTGTGTACCGGCGCGCGCGCGAGCCACCAACGAGCGG 122
QY 121 GAGTGAAGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGCACGTCGGACGG 180
DB 123 GAGTGAAGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGCACGTCGGACGG 182
QY 181 TCCGCGCGCGCGCGCGGAGGAAATGCAATGCGGAGCGGAGCCACCAACGAGCGG 240
DB 183 TCCGCGCGCGCGCGCGGAGGAAATGCAATGCGGAGCGGAGCCACCAACGAGCGG 242
QY 241 CTGGGTGTGAGATATTGAACTGCGGAGCTGTGAGGAGATATTGAAATCAGGATCGG 300
DB 243 CTGGGTGTGAGATATTGAACTGCGGAGCTGTGAGGAGATATTGAAATCAGGATCGG 302
QY 301 CTTAAATTTGAGCCCTTTTATTTGTTGAGCGGTATTCCTGAGTCAAGTTAAAGCTGTT 360
DB 303 CTTAAATTTGAGCCCTTTTATTTGTTGAGCGGTATTCCTGAGTCAAGTTAAAGCTGTT 362
QY 361 GCCGATACAGAAAATATCATGATGCTTAAGGCAACCATGATTTTCAATGTTT 420
DB 363 GCCGATACAGAAAATATCATGATGCTTAAGGCAACCATGATTTTCAATGTTT 422
QY 421 GTGAAGAGAAATGTCAGATGACCTCATTCAGACAGAAATCTATTACCTGGCATGCT 480
DB 423 GTGAAGAGAAATGTCAGATGACCTCATTCAGACAGAAATCTATTACCTGGCATGCT 482
QY 481 GGTGAGAACAGAGAAAATACATGTTTATTTCTGAAATTCCTCAAACTATCAATAGGCA 540
DB 483 GGTGAGAACAGAGAAAATACATGTTTATTTCTGAAATTCCTCAAACTATCAATAGGCA 542
QY 541 GCAGTCTTAAATGCTCTCTTGGAGGCTCTTT 571
DB 543 GCAGTCTTAAATGCTCTCTTGGAGGCTCTTT 573

RESULT 61
A1917735/c
LOCUS
DEFINITION
similar to TR:075273 075273 R26984_1; mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

EST. Homo sapiens (human)

Db 253 TTTTGACCTGTGTAAGTCTCTGGTATACACTGGCTATTAAACCAATGAGGAGTTTA 194
 QY 2923 ATCACAGAAACACAGAAATGATCATCTTTGATACCTGGCCATGTACACNTCTACTC 2982
 Db 193 ATCAACAGAAACACAGAAATGATCATCTTTGATACCTGGCCATGTACACNTCTACTC 134
 QY 2983 CTGAATAATGTTGGTGCATCGAGGGGTCTACGGTTTGGTGTAGTAATCTTAATACCTT 3042
 Db 133 CTGAATAATGTTGGTGCATCGAGGGGTCTACGGTTTGGTGTAGTAATCTTAATACCTT 74
 QY 3043 AACCCCATGCTCAAAATCAAAATGATACATATCTCTGAGAGACCCAGCAGCATACCAAG 3102
 Db 73 AACCCCATGCTCAAAATCAAAATGATACATATCTCTGAGAGACCCAGCAGCATACCAAG 14
 QY 3103 AATTACTTAAAAA 3115
 Db 13 AATTACTTAAAAA 1

RESULT 62
 DA860616
 LOCUS
 DEFINITION
 DA860616 PLACE7 Homo sapiens cDNA clone PLACE7013027 5', mRNA
 sequence.
 ACCESSION
 DA860616
 VERSION
 DA860616.1 GI:82098746
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 563)
 Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
 Yamashita, R., Yamamoto, J., Sekine, M., Teuritani, K., Wakeguri, H.,
 Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
 Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
 Murakawa, K., Iehida, S., Ishibashi, T., Takahashi-Fujii, A.,
 Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
 Diversification of Transcriptional Modulation: Large-scale
 Identification and Characterization of Putative Alternative
 Promoters of Human Genes
 Genome Res. 16 (1), 55-65 (2006)
 16344560
 Contact: Takao Isogai
 FLJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: flj-cdna@nifty.com
 NEDO human cDNA Project (New Energy and Industrial Technology
 Developmental Organization, Japan); cDNA library construction:
 Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
 Research Association for Biotechnology (RAB) and Biotechnology
 Center, National Institute of Technology and Evaluation; 3'-end one
 pass sequencing: RAB.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 COMMENT

FEATURES
 source
 Location/Qualifiers
 1..563
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE7013027"
 /tissue_type="placenta"
 /clone_lib="PLACE7"
 /note="vector: pME18SPL3"
 ORIGIN
 Query Match 18.0%; Score 561; DB 9; Length 563;
 Best Local Similarity 100.0%; Pred. No. 1.3e-291;
 Matches 561; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAGTGCTAAAGCCTCCGAGGCCAAGCGCTGCTACTGCGCGCGCTTCTTAGTGCGG 60

Db 3 AAGTGCTAAAGCCTCCGAGGCCAAGCGCTGCTACTGCGCGCGCTTCTTAGTGCGG 62
 QY 61 GCTTCCGCGCGCTGGGTGTTTCACCGCGCGCGCGCGCGAGGAAGCCACTCAACACGAGGACCG 120
 Db 63 GCTTCCGCGCGCTGGGTGTTTCACCGCGCGCGCGCGCGAGGAAGCCACTCAACACGAGGACCG 122
 QY 121 GAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCGCGCTCCATAGCGCACGTCCGGAGCGG 180
 Db 123 GAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCGCGCTCCATAGCGCACGTCCGGAGCGG 182
 QY 181 TCCGCGCGCGCGCGCGGGAAGGAAAATGCAACATGGCAGCAGCAATGGAAAACAGACAG 240
 Db 183 TCCGCGCGCGCGCGCGGGAAGGAAAATGCAACATGGCAGCAGCAATGGAAAACAGACAG 242
 QY 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300
 Db 243 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG 302
 QY 301 CCTAAATTGGAGCCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAGAGCTGTT 360
 Db 303 CCTAAATTGGAGCCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAGAGCTGTT 362
 QY 361 GCCGATACAGAAAATATCATGCTATGCTAAGGTAAGGCAACCATGATGATTTCAATGTT 420
 Db 363 GCCGATACAGAAAATATCATGCTAAGGTAAGGCAACCATGATGATTTCAATGTT 422
 QY 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTTATACCTTGCATGCT 480
 Db 423 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTTATGCTTGCATGCT 482
 QY 481 GGTGAGAACAGAGAAAATACACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGAGCA 540
 Db 483 GGTGAGAACAGAGAAAATACACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGAGCA 542
 QY 541 GCAGTCTTAATGCTCTCTTGG 561
 Db 543 GCAGTCTTAATGCTCTCTTGG 563

RESULT 63
 CN427235
 LOCUS
 DEFINITION
 CN427235 678 bp mRNA linear EST 16-MAY-2004
 ACCESSION
 CN427235 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
 VERSION
 CN427235.1 GI:47414829
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 678)
 Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
 Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
 Lebkowski, J. and Stanton, L.
 Transcriptional characterization of
 control human ES cell growth and differentiation
 Nat. Biotechnol. 22 (6), 707-716 (2004)
 15146197
 Contact: Brandenberger R
 Regenerative Medicine
 Geron Corporation
 230 Constitution Drive, Menlo Park, CA 94025, USA
 Tel: 650 473 8658
 Fax: 650 473 7766
 Email: rbrandenberger@geron.com
 Insert Length: 678 Std Error: 0.00.
 Location/Qualifiers
 1..678
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 FEATURES
 source
 Location/Qualifiers

Thu Jun 22 09:04:26 2006

Query Match 18.0%; Score 561; DB 8; Length 678;
Best Local Similarity 100.0%; Pred. No. 1.3e-291; Indels 0; Gaps 0;
Matches 561; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/tissue_type="embryonic stem cells, cell lines H1, H7, and H9"
/clone_lib="GRN_ES"
/note="oligo dt primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"
FEATURES
source

ORIGIN
Query Match 17.9%; Score 558; DB 7; Length 673;
Best Local Similarity 99.7%; Pred. No. 5.6e-290; Indels 0; Gaps 0;
Matches 658; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2450 TGATGGCATTAATGCGAGGTCAGATATCTTCAGGGTTCATATGCTGGGGCCCGACCA 2509
DB 660 TGATGGCATTAATGCGAGGTCAGATATCTTCAGGGTTCATATGCTGGGGCCCGACCA 601
QY 2510 CTCGTGGATCTTCTATGATACAGGATACAGGACCTTATATGGGTTCACCTGACCA 2569
DB 600 CTCGTGGATCTTCTATGATACAGGATACAGGACCTTATATGGGTTCACCTGACCA 541
QY 2570 ATGAAAGGGCTTATCTAGGATCTGCGCATGCAAGCAGAGAAAGTTCCTCTGAC 2629
DB 540 ATGAAAGGGCTTATCTAGGATCTGCGCATGCAAGCAGAGAAAGTTCCTCTGAC 481
QY 2630 CAAATCTTCTACCTGCTCTTACATGTTCTCGATGCAATGTCATTTGCAATACCA 2689
DB 480 CAAATCTTCTACCTGCTCTTACATGTTCTCGATGCAATGTCATTTGCAATACCA 421
QY 2690 GTATATCTAGGTTTCTGAGGCTGGAAGCCATATGATTAAGATCTATCTCTC 2749
DB 420 GTATATCTAGGTTTCTGAGGCTGGAAGCCATATGATTAAGATCTATCTCTC 361
QY 2750 AGGAGACACAGCATTAAGAGTTCCTGATCGGAGAACATTTGCAATCTTTTGC 2809
DB 360 AGGAGACACAGCATTAAGAGTTCCTGATCGGAGAACATTTGCAATCTTTTGC 301
QY 2810 ACTACCTTCAAGAAACCTTGGATCAGTATTCCTCTTAAAGTGATATATTTGAC 2869
DB 300 ACTACCTTCAAGAAACCTTGGATCAGTATTCCTCTTAAAGTGATATATTTGAC 241
QY 2870 CTGTGTAGAACTCTCTGTGTATACCTGCTATTTAACCAATGAGGAGTTTAAATCAACA 2929
DB 240 CTGTGTAGAACTCTCTGTGTATACCTGCTATTTAACCAATGAGGAGTTTAAATCAACA 181
QY 2930 GAAAAACACAGATTTGATTCATCTTTGATACCTGCGCATGTAACTCTCTCTGAAA 2989
DB 180 GAAAAACACAGATTTGATTCATCTTTGATACCTGCGCATGTAACTCTCTCTGAAA 121
QY 2990 TAAATGTGGTCCATCAGGGGTCTACCGTTTGTGGTGTATCTTAATACCTTAAACCCCA 3049
DB 120 TAAATGTGGTCCATCAGGGGTCTACCGTTTGTGGTGTATCTTAATACCTTAAACCCCA 61
QY 3050 CATGCTCAAAATCAATGATACATATTCCTGAGAGACCCAGCAATACCATGAATTAAT 3109
DB 60 CATGCTCAAAATCAATGATACATATTCCTGAGAGACCCAGCAATACCATGAATTAAT 1

RESULT 64
AW303607/c
LOCUS
DEFINITION
IMAGE:2813759_3, similar to TR:075273 075273 R26984_1, mRNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
673 bp mRNA linear EST 18-JAN-2000
xv21b12.x1 Soares NFL T GBC S1 Homo sapiens cDNA clone
IMAGE:2813759_3, similar to TR:075273 075273 R26984_1, mRNA
sequence.
AW303607
AW303607.1 GI:6713296
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 673)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.

ORIGIN

Query Match 17.8%; Score 556; DB 9; Length 558;
 Best Local Similarity 100.0%; Pred. No. 6.7e-289;
 Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGTGTAAAGCCTCGAGGCGAAGCGCGCTGCTACTGCGCGCGCTCTTCTAGTGGCG 60
 DB 3 AAGTGTAAAGCCTCGAGGCGCAAGCGCGCTGCTACTGCGCGCGCTCTTCTAGTGGCG 62

QY 61 CGTTCCGCCCTCGGCTGTGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 DB 63 CGTTCCGCCCTCGGCTGTGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 122

QY 121 GAGTGGAGCG 180
 DB 123 GAGTGGAGCG 182

QY 181 TCCGGCG 240
 DB 183 TCCGGCG 242

QY 241 CTGGGCTGTGAGATTTGAACTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
 DB 243 CTGGGCTGTGAGATTTGAACTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 302

QY 301 CCTAAATTCGAGCGCTTTTATGTTGAGCGCGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT 360
 DB 303 CCTAAATTCGAGCGCTTTTATGTTGAGCGCGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT 362

QY 361 GCGGATACAGAAAAATATCATGGCTACATGATGGCTAAGGACACCATGATTTTCATGTTT 420
 DB 363 GCGGATACAGAAAAATATCATGGCTACATGATGGCTAAGGACACCATGATTTTCATGTTT 422

QY 421 GTGAGAGGAATGATCCAGATGGACCTCAATTCAGACAGAACTATTAACCTTGCATGTCT 480
 DB 423 GTGAGAGGAATGATCCAGATGGACCTCAATTCAGACAGAACTATTAACCTTGCATGTCT 482

QY 481 GGTGAGAACAGAGAAAAATACACTGTTTTATCTGAAATTCGAAAACTATCAATAGAGCA 540
 DB 483 GGTGAGAACAGAGAAAAATACACTGTTTTATCTGAAATTCGAAAACTATCAATAGAGCA 542

QY 541 GCAGTCTTAATGCTCT 556
 DB 543 GCAGTCTTAATGCTCT 558

RESULT 67
 BG163397
 LOCUS
 DEFINITION 602338360F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4446375 5', mRNA sequence.

ACCESSION BG163397.1 GI:12670100
 VERSION BG163397.1
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 788)
 NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Plate: LLAM10225 row: g column: 16
 High quality sequence start: 6
 High quality sequence stop: 742.
 Location/Qualifiers
 1. .788
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4446375"
 /tissue_type="hypernephroma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 89"
 /notes="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.3 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 17.8%; Score 555; DB 2; Length 788;
 Best Local Similarity 99.8%; Pred. No. 2.4e-288;
 Matches 605; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2049 TGACTATACCTCCAGAAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGG 2108
 DB 4 TGACTATACCTCCAGAAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGG 63

QY 2109 GATGCTCTACAAGCCTCATGATCTACAGCCTCGAAGAAATATCTACTGTGCTTTCAT 2168
 DB 64 GATGCTCTACAAGCCTCATGATCTACAGCCTCGAAGAAATATCTACTGTGCTTTCAT 123

QY 2169 ATATGTTGTTCCAGTGCGAGTTGTTGAATTAATCGTTTAAAGAGTCAAGTATTTCCG 2228
 DB 124 ATATGTTGTTCCAGTGCGAGTTGTTGAATTAATCGTTTAAAGAGTCAAGTATTTCCG 183

QY 2229 CTTGAATACCTAGCCTCTAGGTTATGTTGTTGATGATAGACAAACAGGGGATCCTG 2288
 DB 184 CTTGAATACCTAGCCTCTAGGTTATGTTGTTGATGATAGACAAACAGGGGATCCTG 243

QY 2289 TCACCGAGGCTTAAATTTGAGGGCGCTTTAAATATAAATGGGTCAANTAGAAATGA 2348
 DB 244 TCACCGAGGCTTAAATTTGAGGGCGCTTTAAATATAAATGGGTCAANTAGAAATGA 303

QY 2349 CGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCG 2408
 DB 304 CGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCG 363

QY 2409 TGTGGGCATCCAGCGTGGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGAG 2468
 DB 364 TGTGGGCATCCAGCGTGGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGAG 423

QY 2469 GTCAGATATCTTCAGGGTGTGCTATTGCTGGGGCCCCCAGTCACCTCTGTGGATCTTCTATGA 2528
 DB 424 GTCAGATATCTTCAGGGTGTGCTATTGCTGGGGCCCCCAGTCACCTCTGTGGATCTTCTATGA 483

QY 2529 TACAGGATACACGGAACGTTTATATGGGTCAACCTGACGGAATGAACAGGGGTATTTACTT 2588
 DB 484 TACAGGATACACGGAACGTTTATATGGGTCAACCTGACGGAATGAACAGGGGTATTTACTT 543

QY 2589 AGGATCTGTGGCCATCAAGCAAGAAAAGTTCCCTCTCTGAACCAAAATCGTTTACTGCTCTT 2648
 DB 544 AGGATCTGTGGCCATCAAGCAAGAAAAGTTCCCTCTCTGAACCAAAATCGTTTACTGCTCTT 603

QY 2649 ACATGG 2654
 DB 604 ACATGG 609

RESULT 68
 DB074943
 LOCUS
 DEFINITION DB074943 TEST14 Homo sapiens cDNA clone TEST14018676 5', mRNA sequence.

DB074943
EST. DB074943.1 GI:83100418
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 553)
Yamashita, R., Yamamoto, J., Sekine, M., Tauritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagaesuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification and Characterization of Putative Alternative Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.
Location/Qualifiers
1. 553
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="F014018676"
/tissue="testis"
/clone_lib="TEST14"
/note="Vector: pME18SFL3"

Query Match
Best Local Similarity 100.0%; Pred. No. 2.8e-287;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

54 AGTGGCGGTTCCGCGCTGGGTGTGCACCGCGCGCGCGAGAGGACCTGCACAC 113
1 AGTGGCGGTTCCGCGCTGGGTGTGCACCGCGCGCGCGAGGACCTGCACAC 60

114 AGACCGGAGTGAGCGCGCGCAGCATGAAGCGCGCGCGCGCTCATAGCGCAGTC 173
61 AGACCGGAGTGAGCGCGCGCAGCATGAAGCGCGCGCGCGCTCATAGCGCAGTC 120

174 GGAGCGGTCGGCGCGCGCGCGGGAAGGAAATGCAATGCGCAGCAGTGGAAAC 233
121 GGAGCGGTCGGCGCGCGCGCGGGAAGGAAATGCAATGCGCAGCAGTGGAAAC 180

234 AGACGAGCTGGGTGTGAGATATTGAACTGCGGACTGTGAGGAGATATTGAATCACA 293
181 AGACGAGCTGGGTGTGAGATATTGAACTGCGGACTGTGAGGAGATATTGAATCACA 240

294 GGAATGGCTAAATTTGGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTCACTTAAAA 353
241 GGAATGGCTAAATTTGGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTCACTTAAAA 300

354 GCTGCTGCCGATACCAAGAAATATCATGCTCATGATGGCTTAAGGACCATGATTT 413
301 GCTGCTGCCGATACCAAGAAATATCATGCTCATGATGGCTTAAGGACCATGATTT 360

414 CATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGATCTTATACCTTGC 473

DB074943
EST. DB074943.1 GI:83100418
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 553)
Yamashita, R., Yamamoto, J., Sekine, M., Tauritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagaesuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification and Characterization of Putative Alternative Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.
Location/Qualifiers
1. 553
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="F014018676"
/tissue="testis"
/clone_lib="TEST14"
/note="Vector: pME18SFL3"

Query Match
Best Local Similarity 100.0%; Pred. No. 2.8e-287;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

54 AGTGGCGGTTCCGCGCTGGGTGTGCACCGCGCGCGCGAGAGGACCTGCACAC 113
1 AGTGGCGGTTCCGCGCTGGGTGTGCACCGCGCGCGCGAGGACCTGCACAC 60

114 AGACCGGAGTGAGCGCGCGCAGCATGAAGCGCGCGCGCGCTCATAGCGCAGTC 173
61 AGACCGGAGTGAGCGCGCGCAGCATGAAGCGCGCGCGCGCTCATAGCGCAGTC 120

174 GGAGCGGTCGGCGCGCGCGCGGGAAGGAAATGCAATGCGCAGCAGTGGAAAC 233
121 GGAGCGGTCGGCGCGCGCGCGGGAAGGAAATGCAATGCGCAGCAGTGGAAAC 180

234 AGACGAGCTGGGTGTGAGATATTGAACTGCGGACTGTGAGGAGATATTGAATCACA 293
181 AGACGAGCTGGGTGTGAGATATTGAACTGCGGACTGTGAGGAGATATTGAATCACA 240

294 GGAATGGCTAAATTTGGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTCACTTAAAA 353
241 GGAATGGCTAAATTTGGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTCACTTAAAA 300

354 GCTGCTGCCGATACCAAGAAATATCATGCTCATGATGGCTTAAGGACCATGATTT 413
301 GCTGCTGCCGATACCAAGAAATATCATGCTCATGATGGCTTAAGGACCATGATTT 360

414 CATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGATCTTATACCTTGC 473

```
QY 81 ACGGGCGCGCGCGGAGGAGCCACTGCAACAGGACCGGAGTGGAGGGCGGCGAGCAT 140
Db 61 ACGGGCGCGCGCGGAGGAGCCACTGCAACAGGACCGGAGTGGAGGGCGGCGAGCAT 120
QY 141 GAAGCGCGCGCGCGCGCTCCATAGCGCACGTCGGGACGGTCCGGGCGGGCGGGGGA 200
Db 121 GAAGCGCGCGCGCGCGCTCCATAGCGCACGTCGGGACGGTCCGGGCGGGCGGGGGA 180
QY 201 AGGAAATGCAACATGCGAGCAGCAATGGAAACAGACAGCTGGGTGAGATATTTCA 260
Db 181 AGGAAATGCAACATGCGAGCAGCAATGGAAACAGACAGCTGGGTGAGATATTTCA 240
QY 261 AACTCGGAGCTGTGAGGAGAAATTTGAATCAACAGATCGGCCTAAATTTGAGGCTTTT 320
Db 241 AACTCGGAGCTGTGAGGAGAAATTTGAATCAACAGATCGGCCTAAATTTGAGGCTTTT 300
QY 321 TGTGAGCGGTATTTCTCGAGTCAGCTTTAAAGCTGCTTGGCGATACCGAATAATCA 380
Db 301 TGTGAGCGGTATTTCTCGAGTCAGCTTTAAAGCTGCTTGGCGATACCGAATAATCA 360
QY 381 TGGCTACATGATGGCTTAAGGACCAACATGATTTTCAATGTTGTAAGAGGAATGATCCAGA 440
Db 361 TGGCTACATGATGGCTTAAGGACCAACATGATTTTCAATGTTGTAAGAGGAATGATCCAGA 420
QY 441 TGGACCTCATTCAGACAGAACTATTACCTTGGCATGTCTGGTGAGAAACAGAGAAATAC 500
Db 421 TGGACCTCATTCAGACAGAACTATTACCTTGGCATGTCTGGTGAGAAACAGAGAAATAC 480
QY 501 ACTGTTTTATCTGAAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTG 560
Db 481 ACTGTTTTATCTGAAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTG 540
QY 561 GAAGCCTCTTTT 572
Db 541 GAAGCCTCTTTT 552
```

```
RESULT 70
CN427230
LOCUS CN427230 553 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000532204093 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN427230
VERSION CN427230.1 GI:47414824
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

```
REFERENCE
AUTHORS Brandenberger R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
TITLE Transcriptional characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R
Regenerative Medicine
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 553 Std Error: 0.00.
```

```
FEATURES
source
1..553
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, cell lines H1, H7, and
H9"
/clone_lib="GRN_ES"
```

```
/note="oligo dT primed, full-length enriched cDNA library
from undifferentiated hES cell lines H1 (p32), H7 (p29),
and H9 (p26) maintained in feeder-free conditions"
```

ORIGIN

```
Query Match 17.6%; Score 548; DB 8; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.5e-284;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 24 AGGCCGCTGCTACTGCGCGCTGCTTCTTAGTGGCGGCTTCGCCGCTTGGGTTGTCAAC 83
Db 6 AGGCCGCTGCTACTGCGCGCTGCTTCTTAGTGGCGGCTTCGCCGCTTGGGTTGTCAAC 65
QY 84 GCGCGCGCGCGCGGAGGACCACTGCAACAGGACCGGAGTGGAGGGCGGCGAGCATGAA 143
Db 66 GCGCGCGCGCGCGGAGGAGCCACTGCAACAGGACCGGAGTGGAGGGCGGCGAGCATGAA 125
QY 144 GCGCGGAGGCGCGCTTCATAGCGCACGTCGGGACGGTCCGGGCGGGGCGGGGGAAGG 203
Db 126 GCGCGCGAGCGCGCTCCATAGCGCACGTCGGGACGGTCCGGGCGGGGCGGGGGAAGG 185
QY 204 AAAATGCAACATGCGGAGCAATGCAACAGGACCGGAGTGGAGATATTTGAAAC 263
Db 186 AAAATGCAACATGCGGAGCAATGCAACAGGACCGGAGTGGAGATATTTGAAAC 245
QY 264 TCGGAGCTGTGAGGAGAAATTTGAATCAAGGATCGGCCTAAATTTGAGACCTTTTATGT 323
Db 246 TCGGAGCTGTGAGGAGAAATTTGAATCAAGGATCGGCCTAAATTTGAGACCTTTTATGT 305
QY 324 TGAGCGGTATCTCGAGTCAGCTTTAAAGCTGCTTCCCGATACCGAATAATCATGG 383
Db 306 TGAGCGGTATCTCGAGTCAGCTTTAAAGCTGCTTCCCGATACCGAATAATCATGG 365
QY 384 CTACATGATGGCTAAGGCACCAATGATTTTCAATGTTGTGAAGAGGAATGATCCAGATGG 443
Db 366 CTACATGATGGCTAAGGCACCAATGATTTTCAATGTTGTGAAGAGGAATGATCCAGATGG 425
QY 444 ACCTCATTCAGACAGAACTATTACCTTGGCATGTCTGGTGAGAAACAGAGAAATACACT 503
Db 426 ACCTCATTCAGACAGAACTATTACCTTGGCATGTCTGGTGAGAAACAGAGAAATACACT 485
QY 504 GTTTTATCTGAAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTGGAA 563
Db 486 GTTTTATCTGAAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTGGAA 545
QY 564 GCCTCTTT 571
Db 546 GCCTCTTT 553
```

```
RESULT 71
CN427227 612 bp mRNA linear EST 16-MAY-2004
LOCUS CN427227
DEFINITION 17000532648202 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN427227
VERSION CN427227.1 GI:47414821
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

```
REFERENCE
AUTHORS Brandenberger R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
TITLE Transcriptional characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
```

230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 612 Std Error: 0.00.
Location/Qualifiers

FEATURES
source

1. .612
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, cell lines H1, H7, and H9"
/clone_lib="GRN ES"
/note="oligo at_1 primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

ORIGIN

Query Match 17.6%; Score 548; DB 8; Length 612;
Best Local Similarity 99.8%; Pred. No. 1.5e-284;
Matches 598; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 24 AGGCGCTGCTACTGCGCGCTGCTTCTTAGTGGCGCTTCGCGCGCTGGTGTGTACCC 83
Db |||||
QY 14 AGGCGCTGCTACTGCGCGCTGCTTCTTAGTGGCGCTTCGCGCGCTGGTGTGTACCC 73
Db |||||
QY 84 GGCGCGCGCGCGGAGGAGCACTGCAACAGGACCGGAGTGGAGCGCGCGCGCATGAA 143
Db |||||
QY 74 GGCGCGCGCGCGGAGGAGCACTGCAACAGGACCGGAGTGGAGCGCGCGCATGAA 133
Db |||||
QY 144 GGCGCGCGCGCGCTCCATAGCGCACGTCGGGACGCTCCGGCGCGCGCGGGGAGG 203
Db |||||
QY 134 GGCGCGCGCGCGCTCCATAGCGCACGTCGGGACGCTCCGGCGCGCGCGGGGAGG 193
Db |||||
QY 204 AAAATGCAATCGCAGCAGCAATGGAACAGAACAGTGGTGTGAGTATTTGAAAC 263
Db |||||
QY 194 AAAATGCAATCGCAGCAGCAATGGAACAGAACAGTGGTGTGAGTATTTGAAAC 253
Db |||||
QY 264 TGGCGACTGTGAGGAGATTTGAATCACAGATCGGCTAAATTTGAGCGCTTTTATGT 323
Db |||||
QY 254 TGGCGACTGTGAGGAGATTTGAATCACAGATCGGCTAAATTTGAGCGCTTTTATGT 313
Db |||||
QY 324 TGAGCGGTATTCCTGGAGTCAGCTTAAAGCTCTCTCCGATACAGAGAAATATCATGG 383
Db |||||
QY 314 TGAGCGGTATTCCTGGAGTCAGCTTAAAGCTCTCTCCGATACAGAGAAATATCATGG 373
Db |||||
QY 384 CTACATGATGGCTAAGGACACATGATTTTCATGCTTTGTGAAGGAGTATTCAGATGG 443
Db |||||
QY 374 CTACATGATGGCTAAGGACACATGATTTTCATGCTTTGTGAAGGAGTATTCAGATGG 433
Db |||||
QY 444 ACCTATTTCAGACAGAACTATTACCTTTGCCATGCTGTGAGAACAGAGAAATACACT 503
Db |||||
QY 434 ACCTATTTCAGACAGAACTATTACCTTTGCCATGCTGTGAGAACAGAGAAATACACT 493
Db |||||
QY 504 GTTTTATTCTGAAATTCCTCAAACTATCAATAGAGCGAGCTTAATGCTCTCTTGAA 563
Db |||||
QY 494 GTTTTATTCTGAAATTCCTCAAACTATCAATAGAGCGAGCTTAATGCTCTCTTGAA 553
Db |||||
QY 564 GCCTCTTTTGGATCTTTTCAGGCACTGGAATGGAATGTATTCGAGAGAGAG 622
Db |||||
QY 554 GCCTCTTTTGGATCTTTTCAGGCACTGGAATGGAATGTATTCGAGAGAGAG 612
Db |||||

RESULT 72
AF173382
LOCUS AF173382 1278 bp mRNA linear HTC 01-AUG-2003
DEFINITION Homo sapiens MSTP097 (MST097) mRNA, complete cds.
ACCESSION AF173382
VERSION AF173382.1 GI:33338001
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo (1278)
1 (bases 1 to 1278)
Xu, H. S. Zhao, B., Tong, Y. K., Sheng, H., Qin, B. M., Liu, Y. Q., Liu, B.,
Wang, X. Y., Zhang, Q., Song, L., Gao, Y., Zhang, C. L., Ye, J., Ji, X. J.,
Liu, B. H., Lu, H., Chen, J. Z., Cai, M. Q., Zheng, W. Y., Teng, C. Y.,
Liu, Q., Yu, L. T., Lin, J., Gong, Q., Zhang, A. M., Gao, R. L. and Hiu, R. T.
Homo sapiens normal aorta mRNA MST097
Unpublished
2 (bases 1 to 1278)
Xu, H. S. Zhao, B., Tong, Y. K., Sheng, H., Qin, B. M., Liu, Y. Q., Liu, B.,
Wang, X. Y., Zhang, Q., Song, L., Gao, Y., Zhang, C. L., Ye, J., Ji, X. J.,
Liu, B. H., Lu, H., Chen, J. Z., Cai, M. Q., Zheng, W. Y., Teng, C. Y.,
Liu, Q., Yu, L. T., Lin, J., Gong, Q., Zhang, A. M., Gao, R. L. and Hiu, R. T.
Direct Submission
Submitted (30-JUL-1999) Molecular Medicine Center for
Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167
Bei Li Shi Lu, Beijing 100037, P.R. China

FEATURES
source

1. .1278
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="aorta"
1. .1278
/genes="MST097"
591..1010
/genes="MST097"
/codon_start=1
/product="MSTP097"
/protein_id="AAQ13623.1"
/db_xref="GI:33338002"
/translation="MEIPSPDGICRGIOISSGAIAGPSHSVISNQHGRVGHLMTRN
QGYVLGSVAMQAEKFPNRLLLHLGFLDENVHPATISILLSLFLVRAGPKYDQIYP
QERHSIRVPESEGHVELLHLHLQENLGSRIALKVI"

ORIGIN

Query Match 17.5%; Score 547; DB 6; Length 1278;
Best Local Similarity 100.0%; Pred. No. 5.4e-284;
Matches 547; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1844 CTGGAGAGTGCACAGGCTGACTGACGCTGGCTACTCATTCTTCTGTCATCAGTCAGC 1903
Db |||||
QY 1 1 CTGGAGAGTGCACAGGCTGACTGACGCTGGCTACTCATTCTTCTGTCATCAGTCAGC 60
Db |||||
QY 1904 ACTGTGACTTCTTTATAGTAAGTATAGTACCGAGAGATCCACACTGTGTCTCCTTT 1963
Db |||||
QY 61 ACTGTGACTTCTTTATAGTAAGTATAGTACCGAGAGATCCACACTGTGTCTCCTTT 120
Db |||||
QY 1964 ACAAGCTATCAAGCTCTGAAGTGCACCACTTGCACCAAAAGAAATTTGGGCCACCA 2023
Db |||||
QY 121 ACAAGCTATCAAGCTCTGAAGTGCACCACTTGCACCAAAAGAAATTTGGGCCACCA 180
Db |||||
QY 2024 TTTTGGATTCAGAGGCTCTTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTGAAA 2083
Db |||||
QY 181 TTTTGGATTCAGAGGCTCTTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTGAAA 240
Db |||||
QY 2084 GTACTACTGGATTTACATTTGTATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGA 2143
Db |||||
QY 241 GTACTACTGGATTTACATTTGTATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGA 300
Db |||||
QY 2144 AGAAATATCTTACTGTGCTCTTATATATGGTGGTCTCAGGTGCAGTTGGTGAATAATC 2203
Db |||||
QY 301 AGAAATATCTTACTGTGCTCTTATATATGGTGGTCTCAGGTGCAGTTGGTGAATAATC 360
Db |||||
QY 2204 GGTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGGTTG 2263
Db |||||
QY 361 GGTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGGTTG 420
Db |||||
QY 2264 TAGTGATAGACAAAGGGGATCTGTCTACCGAGGCTTAAATTTGAAGGCCCTTTAAAT 2323
Db |||||
QY 421 TAGTGATAGACAAAGGGGATCTGTCTACCGAGGCTTAAATTTGAAGGCCCTTTAAAT 480
Db |||||


```
QY 2324 ATAAATGGGTCAAAATAGAAATTCAGATCAGGTGGAAGCACTCCAAATATCTAGCTTCTC 2383
|||||
Db 481 ATAAATGGGTCAAAATAGAAATTCAGATCAGGTGGAAGCACTCCAAATATCTAGCTTCTC 540
|||||
QY 2384 GATATGA 2390
|||||
Db 541 GATATGA 547
|||||

RESULT 73
DA052731 546 bp mRNA linear EST 30-OCT-2005
LOCUS DA052731 BNGH42 Homo sapiens cDNA clone BNGH42008161 5', mRNA
DEFINITION
ACCESSION DA052731
VERSION DA052731.1 GI:78278014
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 546)
AUTHORS Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kishida,N.,
Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
Murakawa,K., Ishida,S., Iehibashi,T., Takahashi-Fujii,A.,
Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO Human CDNA Project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
FEATURES
    source
        1..546
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="BNGH42008161"
            /cell_type="neuroglioma"
            /cell_line="H4"
            /clone_lib="BNGH42"
            /note="Vector: pME18SFL3"

ORIGIN
Query Match 17.5%; Score 546; DB 9; Length 546;
Best Local Similarity 100.0%; Pred. No. 1.8e-283; Indels 0; Gaps 0;
Matches 546; Conservative 0; Mismatches 0;

QY 51 CTTAGTCCGCGTTCGCGCCTGGGTGTCTACCGCGCGCGCGCGGAGCACTGCA 110
Db 1 CTTAGTCCGCGTTCGCGCCTGGGTGTCTACCGCGCGCGCGCGGAGCACTGCA 60
|||||
QY 111 ACCAGACCGGAGTGGAGCGCGCGAGCATGACGGCGGAGCGCCCTCATAGCGCAC 170
Db 61 ACCAGACCGGAGTGGAGCGCGCGAGCATGACGGCGGAGCGCCCTCATAGCGCAC 120
|||||
QY 171 GTCGGACCGGTCCGGCGCGGCGCGGGGGAAGAAATGCAACATGCGACGCAATGGA 230
|||||
```

(University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN	Query Match	Best Local Similarity	17.4%; Score 543; DB 3; Length 897;	Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1635	ATATAACGATCCAGTGGTGGCTGCTGCTCCAGTGAATTCAGGTGCTCTATCAAGA	1694	
Db	1	ATATAACGATCCAGTGGTGGCTGCTGCTCCAGTGAATTCAGGTGCTCTATCAAGA	60	
Qy	1695	GGAGATAGCAATACCAAGTGGTGAATGGGAAGTTCCTGGCGGCGATGGATCTAATATCA	1754	
Db	61	GGAGATAGCAATACCAAGTGGTGAATGGGAAGTTCCTGGCGGCGATGGATCTAATATCA	120	
Qy	1755	AGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCA	1814	
Db	121	AGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCA	180	
Qy	1815	TCACCTGTACGTAGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT	1874	
Db	181	TCACCTGTACGTAGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT	240	
Qy	1875	CTACTCACATCTTCTGCTGCATCAGTCAGTACGCTGCTCTCTTTAATGAAGTATAGTAA	1934	
Db	241	CTACTCACATCTTCTGCTGCATCAGTCAGTACGCTGCTCTCTTTAATGAAGTATAGTAA	300	
Qy	1935	CCAGAGAATCCACACTGTGTGCTCTTTTACAGCTATCAAGTCTCTGAAGATGACCCAAAC	1994	
Db	301	CCAGAGAATCCACACTGTGTGCTCTTTTACAGCTATCAAGTCTCTGAAGATGACCCAAAC	360	
Qy	1995	TTGCAAAACAAAGGAATTTTGGCCACCAATTTTGGATTTCAGCAGGCTCTCTTCTGACTA	2054	
Db	361	TTGCAAAACAAAGGAATTTTGGCCACCAATTTTGGATTTCAGCAGGCTCTCTTCTGACTA	420	
Qy	2055	TACTCTCCAGAAATTTCTCTTTGAAGTACTACTGGAATTCATGATGATGGATGCT	2114	
Db	421	TACTCTCCAGAAATTTCTCTTTGAAGTACTACTGGAATTCATGATGATGGATGCT	480	
Qy	2115	CTACAGCTCATGATCTACAGCTTGGAAAGAAATATCTACTGTGCTGTTTATATATGG	2174	
Db	481	CTACAGCTCATGATCTACAGCTTGGAAAGAAATATCTACTGTGCTGTTTATATATGG	540	
Qy	2175	TGG 2177		
Db	541	TGG 543		

RESULT 75
 DA809185 LOCUS DA809185 541 bp mRNA linear EST 11-NOV-2005
 DEFINITION DA809185 OCBFP3 Homo sapiens cDNA clone OCBFP3025332 5', mRNA
 ACCESSION DA809185
 VERSION DA809185.1 GI:82115632
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 541)
 Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Teurittani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, I., and Sugano, S.
 Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes

JOURNAL
 PUBMED
 COMMENT

Genome Res. 16 (1), 55-65 (2006)

Contact: Takao Isogai
 FLJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa, Kametani, Kisarazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3966
 Email: flj_cdn@hri.com

NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction; Helix Research Institute (HRI); 5'-end one pass sequencing; HRI, Center, National Institute of Technology (RAB) and Biotechnology, pass sequencing; RAB.

FEATURES
 source

Location/Qualifiers
 1..541
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="OCBFP3025332"
 /tissue_type="brain"
 /dev_stage="fetal"
 /clone_lib="OCBFP3"
 /note="vector: pME18SFL3"

ORIGIN

Query Match 17.3%; Score 541; DB 9; Length 541;
 Best Local Similarity 100.0%; Pred. No. 9,1e-281; Indels 0; Gaps 0;
 Matches 541; Conservative 0; Mismatches 0;

Qy	2	AGTGCTAAAGCCCTCCGAGGCCAAGCGCTGCTACTGCGCGCGCTGCTTCTAGTGGCGC	61	
Db	1	AGTGCTAAAGCCCTCCGAGGCCAAGCGCTGCTACTGCGCGCGCTGCTTCTAGTGGCGC	60	
Qy	62	GTTGCGCGCTGCTGGTGTGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	121	
Db	61	GTTGCGCGCTGCTGGTGTGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	120	
Qy	122	AGTGGAGGGCG	181	
Db	121	AGTGGAGGGCG	180	
Qy	182	CCGGCG	241	
Db	181	CCGGCG	240	
Qy	242	TGGGTGTTGAGATATTTGAACTCGGACTGTGAGGAGAAATATGAATCAAGGATCGGC	301	
Db	241	TGGGTGTTGAGATATTTGAACTCGGACTGTGAGGAGAAATATGAATCAAGGATCGGC	300	
Qy	302	CTAAATTCGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTTG	361	
Db	301	CTAAATTCGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTTG	360	
Qy	362	CCGATACAGAAATATCATGCTCATGATGCTAAGGCACCAATGATGATTTTCTGTTTG	421	
Db	361	CCGATACAGAAATATCATGCTCATGATGCTAAGGCACCAATGATGATTTTCTGTTTG	420	
Qy	422	TGAAGAGATGATCCAGATGGAACCTCATTCAGACAGAAATCTATTACCTTGCATGCTTG	481	
Db	421	TGAAGAGATGATCCAGATGGAACCTCATTCAGACAGAAATCTATTACCTTGCATGCTTG	480	
Qy	482	GTGAGACAGAGAAATACACTGTTTATCTGAAATTCCTCCAAATCTCAATAGAGCAG	541	
Db	481	GTGAGACAGAGAAATACACTGTTTATCTGAAATTCCTCCAAATCTCAATAGAGCAG	540	
Qy	542	C 542		
Db	541	C 541		

RESULT 76

DB339434/C
LOCUS DB339434 TEST14 Homo sapiens cDNA clone linear EST 04-DEC-2005
DEFINITION DB339434 sequence.
ACCESSION DB339434
VERSION DB339434.1 GI:83174333
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
AUTHORS Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushiida,N.,
Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
TITLE Diversification of transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
JOURNAL Genome Res. 16 (1), 55-65 (2006)
PUBMED 16344560
COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEBO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
FEATURES
source Location/Qualifiers
1..556
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TEST14003235"
/tissue_type="testis"
/clone_lib="TEST14"
/note="Vector: pME18SFL3"
ORIGIN
Query Match 17.3%; Score 541; DB 9; Length 556;
Best Local Similarity 100.0%; Pred. No. 9.2e-281; Indels 0; Gaps 0;
Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2559 CCCTGACAGATGAACAGGCGCTATTACTTAGGATCTGTGGCCATCAAGCAGAAAGTT 2618
DB 541 CCCTGACAGATGAACAGGCGCTATTACTTAGGATCTGTGGCCATCAAGCAGAAAGTT 482
QY 2619 CCCCTCTGAACCAATCGTTTACTGTCTTACATGGTTTCTCGATGAGAATGTCATT 2678
DB 481 CCCCTCTGAACCAATCGTTTACTGTCTTACATGGTTTCTCGATGAGAATGTCATT 422
QY 2679 TGCACATACAGTATTACTGAGTGTGTTTTAGTGAGGGCTGGAAAGCCATATGATTACA 2738
DB 421 TGCACATACAGTATTACTGAGTGTGTTTTAGTGAGGGCTGGAAAGCCATATGATTACA 362
QY 2739 GATCTATCTCTCAGGAGAGACACAGCATAGAGTTCTGATCGGGAGAACATTTCAACT 2798
DB 361 GATCTATCTCTCAGGAGAGACACAGCATAGAGTTCTGATCGGGAGAACATTTCAACT 302
QY 2799 GCATCTTTTGCATCTTCAAGAAACCTTGGATCAGTATTGCTGCTCTTAAAGTGAT 2858
DB 301 GCATCTTTTGCATCTTCAAGAAACCTTGGATCAGTATTGCTGCTCTTAAAGTGAT 242
QY 2859 ATATATTTTGACCTGTGTAGAACTCTCTGATATACACTGGCTATTATTAACCAATGAGGAGG 2918

DB 241 ATAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTATTAACCAATGAGGAGG 182
QY 2919 TTTAATCAACAGAAAACACAGAAATTTGATCATCATCATTTTGTGATACCTGCCATGTAACTCT 2978
DB 181 TTTAATCAACAGAAAACACAGAAATTTGATCATCATCATTTTGTGATACCTGCCATGTAACTCT 122
QY 2979 ACTCTCTGAAAATAAATGTGTGCCATGCGGGGCTCTACGGTTTGTGTAGTAGTAACTTAATA 3038
DB 121 ACTCTCTGAAAATAAATGTGTGCCATGCGGGGCTCTACGGTTTGTGTAGTAGTAACTTAATA 62
QY 3039 CTTTAACCCCATGCTCTCAAAATCAATGATACATATTTCTGAGAGACCCAGCAATACCA 3098
DB 61 CTTTAACCCCATGCTCTCAAAATCAATGATACATATTTCTGAGAGACCCAGCAATACCA 2
QY 3099 T 3099 540 bp mRNA linear EST 06-NOV-2005
DB DA437440 CTONG2 Homo sapiens cDNA clone CTONG2004782 5', mRNA
LOCUS DA437440
DEFINITION DA437440 sequence.
ACCESSION DA437440
VERSION DA437440.1 GI:80893393
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 540)
AUTHORS Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushiida,N.,
Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
TITLE Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
JOURNAL Genome Res. 16 (1), 55-65 (2006)
PUBMED 16344560
COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEBO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
FEATURES
source Location/Qualifiers
1..540
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CTONG2004782"
/tissue_type="tongue, tumor tissue"
/clone_lib="CTONG2"
/note="Vector: pME18SFL3"
ORIGIN
Query Match 17.3%; Score 539; DB 9; Length 540;
Best Local Similarity 100.0%; Pred. No. 1.1e-279; Indels 0; Gaps 0;
Matches 539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.

Location/Qualifiers
1. .569
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TRACH2007313"
/tissue_type="trachea"
/clone_lib="TRACH2"
/note="Vector: pME18SF13"

FEATURES
source

Query Match	17.3%;	Score 539;	DB 9;	Length 569;
Best Local Similarity	100.0%;	Pred. No. 1.1e-279;	Indels 0;	Gaps 0;
Matches 539;	Conservative 0;	Mismatches 0;		
QY	1	AAGTGTAAAGCCTCCGAGGCCAAGGCGCGTCTACTATGCGCGCGCTGCTTCTTAGTGCGG	60	
DB	1	AAGTGTAAAGCCTCCGAGGCCAAGGCGCGTCTACTATGCGCGCGCTGCTTCTTAGTGCGG	60	
QY	61	CGTTGCGCGCGCTGGTGTTCACCGGCGCGCGCGCGCGAGGAGCCACTGTCAACACAGACCG	120	
DB	61	CGTTGCGCGCGCTGGTGTTCACCGGCGCGCGCGCGAGGAGCCACTGTCAACACAGACCG	120	
QY	121	GAGTGGAGGCGCGCAGCATTGAAGCGCGCAGGCGCGCTCCATAGCGCAGCTCGGGACGG	180	
DB	121	GAGTGGAGGCGCGCAGCATTGAAGCGCGCGCAGGCGCGCTCCATAGCGCAGCTCGGGACGG	180	
QY	181	TCCGGGCGCGCGCCGGGGGGAAGGAAATGCACATGGCGACGCMATGGAAACAGACAG	240	
DB	181	TCCGGGCGCGCGCCGGGGGGAAGGAAATGCACATGGCGACGCMATGGAAACAGACAG	240	
QY	241	CTGGGTGTGAGATATTTGAAACTGCGGCACTGTGAGGAGAAATATTGAAATCAACGATCGG	300	
DB	241	CTGGGTGTGAGATATTTGAAACTGCGGCACTGTGAGGAGAAATATTGAAATCAACGATCGG	300	
QY	301	CCTAAATTTGGAGCCCTTTTATGTGTGACGGGTATCTCTGGAGTCAGCTTAAAAAAGCTGCTT	360	
DB	301	CCTAAATTTGGAGCCCTTTTATGTGTGACGGGTATCTCTGGAGTCAGCTTAAAAAAGCTGCTT	360	
QY	361	GCCGATACCAAGAAATATCATGGCTACATGATGGCTAAGGCACCAATGATTTCAATGTTT	420	
DB	361	GCCGATACCAAGAAATATCATGGCTACATGATGGCTAAGGCACCAATGATTTCAATGTTT	420	
QY	421	GTGAAGAGGAATGATCCAGATGGACCTCATTACAGAGAAATCTATTACCTTGCCTATGCT	480	
DB	421	GTGAAGAGGAATGATCCAGATGGACCTCATTACAGAGAAATCTATTACCTTGCCTATGCT	480	
QY	481	GGTGAGAAACAGAGAAATACACTGTTTTTATTCTGAAATTCGAAAACCTATCAATAGAC	539	
DB	481	GGTGAGAAACAGAGAAATACACTGTTTTTATTCTGAAATTCGAAAACCTATCAATAGAC	539	

ORIGIN

RESULT 80

DA446578

RESULT 80
DA446578

LOCUS	DA446578	591 bp	mrna	linear	EST 06-NOV-2005
DEFINITION	DA446578 CTONG2 Homo sapiens cDNA clone CTONG2017644 5', mRNA sequence.				
ACCESSION	DA446578				
VERSION	DA446578.1	GI:80962067			
KEYWORDS	EST.				
SOURCE	Hominidae; Homo.				
ORGANISM	Homo sapiens (human)				
REFERENCE	1 (bases 1 to 591)				
AUTHORS	Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaquri, H., Iehi, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Murakawa, I., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.				
TITLE	Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes				
JOURNAL	Genome Res. 16 (1), 55-65 (2006)				
PUBMED	16344560				
COMMENT	Contact: Takao Isogai FLJ Project (HRI Team) Helix Research Institute 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: flj-cdna@nifty.com NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.				
FEATURES	1..591				
source	1..591				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="CTONG2017644"				
	/issue_type="congue, tumor tissue"				
	/clone_lib="CTONG2"				
	/note="Vector: pME18SFL3"				

Query Match	17.3%	Score 539	DB 9	Length 591
Best Local Similarity	99.8%	Pred. No. 1.1e-279		
Matches 589	Conservative 0	Mismatches 1	Indels 0	Gaps 0
Qy	506	TTTATTCTGAAATTC	CCAAAATCAATAGAGACAGCTCTTAATGCTCTCTTCGAAGC	565
Db	2	TTTATCTGAAATTC	CCAAAATCAATAGAGCGCGCTTAAATGCTCTCTTCGAAGC	61
Qy	566	CTCTTTTGGATCTTTT	CAGCGCAACCTGGACTATGGAATGTATTCTCGAGAAGAAGAAC	625
Db	62	CTCTTTTGGATCTTTT	CAGGCAACACTGGGACTATGGAATGTATTCTCGAGAAGAAGAAC	121
Qy	626	TATTAAAGAGAAAGAAA	CGCATTTGCGAATTCCTTACGATTATCACAAG	685
Db	122	TATTAAAGAGAGAAA	CGCATTTGCGAATTCCTTACGATTATCACAAG	181
Qy	686	GAAGTGGAAACATTTCTGTTT	CAAGCCGTAGTGGAAATTTATCAGTAAAAAGATGGAGGC	745
Db	182	GAAGTGGAAACATTTCTGTTT	CAAGCCGTAGTGGAAATTTATCAGTAAAAAGATGGAGGC	241
Qy	746	CACAAGGATTTACGCAACAACTTT	TAAGGCCCAATCTAGTGGAAACCTAGTTGTCCCAACA	805
Db	242	CACAAGGATTTACGCAACAACTTT	TAAGGCCCAATCTAGTGGAAACCTAGTTGTCCCAACA	301
Qy	806	TACGGATGGATCCAAAATTA	TGCCCCGTGATCGAGCTGGATTGCTTTTATACATAGCA	865

302 TACGATGGATCCAAATATATGCTGCTGATCCAGACTGATTTTATATACATGCA 361
866 ACATATTTGGATATTAACATCGTAACACAGAGAAGAGAGACTCATTATGTGACA 925
362 ACATATTTGGATATTAACATCGTAACACAGAGAAGAGAGACTCATTATGTGACA 421
926 ATGAGCTAGCCACATGGAAGAAGATGCGATCAGCTGAGTGGCTACCTTTGTTCTCC 985
422 ATGAGCTAGCCACATGGAAGAAGATGCGATCAGCTGAGTGGCTACCTTTGTTCTCC 481
986 AAGAAGAATTTGATAGATATTTCTGGCTATTTGCTGCTCCAAAGCTGAAACATCCCA 1045
482 AAGAAGAATTTGATAGATATTTCTGGCTATTTGCTGCTCCAAAGCTGAAACATCCCA 541
1046 GTGGTGGTAAAAATTTCTAGATTTCTATATGAGAAATGATGATCTGAG 1095
542 GTGGTGGTAAAAATTTCTAGATTTCTATATGAGAAATGATGATCTGAG 591

RESULT 81
BU784521/c
LOCUS
DEFINITION
in6d12.x1 Human insulinoma Homo sapiens cdna clone IMAGE:6124487
3' similar to TR:075273 OT5273 R26984_1, mRNA sequence.
ACCESSION
BU784521
VERSION
BU784521.1 GI:23829606
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo
1 (bases 1 to 597)
Melton, D. Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemihela, I., Secombe, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hiller, K., Marra, M., Pape, D., Wylie, T., Martin, J., Blissett, A.,
Schmidt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Teagarsiehl, R.,
Williams, T., Jackson, Y., and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. J. Ferrer In vivo mass-excised to
pBlueScript SK- by Dr. H. Inoue DNA sequencing by: Washington
University Genome Sequencing Center For information on obtaining a
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40UP from Gibco
High quality sequence stop: 490.
Location/Qualifiers
1..597
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6124487"
/tissue="pancreas"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Human insulinoma"
/note="Organ: pancreas; Vector: pBlueScript SK-; Site 1:
XhoI; Site 2: EcoRI; Constructed with lambda ZAPRII system
(Stratagene) by Dr. J. Ferrer in vivo mass-excised to
pBlueScript SK- by Dr. H. Inoue following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Hiroshi Inoue MD/PhD for further
information on this library (Metabolism Division, Permutt
Laboratory, Washington University School of Medicine, Box

8127, 660 S Euclid Ave, St. Louis, MO 63110. Note: this
is a Washington University Pancreas EST project library. "

Query Match 17.2%; Score 538; DB 3; Length 597;
Best Local Similarity 99.8%; Fred. No. 3.9e-279;
Matches 588; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2532 AGGATACACGGAACCTTATATGGGTCCCTGACCAAGTGAACAGGGCTATTACTTAGG 2591
DB 593 AGGATACACGGAACCTTATATGGGTCCCTGACCAAGTGAACAGGGCTATTACTTAGG 534
QY 2592 ATCTGTGGCCATGCGACGAGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTTTACA 2651
DB 533 ATCTGTGGCCATGCGACGAGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTTTACA 474
QY 2652 TCGTTTCTGGATGAGATGTCATTTTGGACATACCAATGATATTAATGAGTTTTTACT 2711
DB 473 TCGTTTCTGGATGAGATGTCATTTTGGACATACCAATGATATTAATGAGTTTTTACT 414
QY 2712 GAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAGAGT 2771
DB 413 GAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAGAGT 354
QY 2772 TCCTGAATCGGAGAACCATTTAATGAACTGCACTCTTTTGCACATCTTTTGCACATCTCTGGTATA 2831
DB 353 TCCTGAATCGGAGAACCATTTAATGAACTGCACTCTTTTGCACATCTTTTGCACATCTCTGGTATA 294
QY 2832 ATACGCTATTTGCTGCTCTAAAAGTGATATAATTTTGAATGCACTCTTTTGCACATCTCTGGTATA 2891
DB 293 ATACGCTATTTGCTGCTCTAAAAGTGATATAATTTTGAATGCACTCTTTTGCACATCTCTGGTATA 234
QY 2892 CACTGGCTATTTAAACCAAAATGAGGAGTTTAAATCAACAGAAAACACAGATGATGATCA 2951
DB 233 CACTGGCTATTTAAACCAAAATGAGGAGTTTAAATCAACAGAAAACACAGATGATGATCA 174
QY 2952 CATTTTGATACCTGCGCATGTAACATCTACTCTCTGAAAATTAATGTCGTCATCGAGGG 3011
DB 173 CATTTTGATACCTGCGCATGTAACATCTACTCTCTGAAAATTAATGTCGTCATCGAGGG 114
QY 3012 TCTACGGTTTGTGGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAATCAATGATAC 3071
DB 113 TCTACGGTTTGTGGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAATCAATGATAC 54
QY 3072 ATATTCTCAGAGACCCAGCAATACCAATGATGATGATGATGATGATGATGATGATGAT 3120
DB 53 ATATTCTCAGAGACCCAGCAATACCAATGATGATGATGATGATGATGATGATGATGAT 5

DA515446 534 bp mRNA linear EST 06-NOV-2005
DA515446 FEBR2 Homo sapiens cdna clone FEBR2003532 5', mRNA
sequence.
ACCESSION
DA515446
VERSION
DA515446.1 GI:80887613
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo
1 (bases 1 to 534)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yanashita, R., Yamamoto, J., Sekine, M., Teuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N.,
Yonekawa, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Tanaka, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T., and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)

RESULT 82
DA515446
LOCUS
DEFINITION
DA515446.1
ACCESSION
DA515446
VERSION
DA515446.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo
1 (bases 1 to 534)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yanashita, R., Yamamoto, J., Sekine, M., Teuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N.,
Yonekawa, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Tanaka, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T., and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)

83

[illegible]

TITLE

FEATURES

ORIGIN

```

Query Match      17.1%; Score 534; DB 9; Length 585;
Best Local Similarity 99.8%; Pred. No. 5.8e-277;
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY	54	AGTCCGCGTTTCGCGCGCTCGGTTGTCAACGGCGCGCCGCTCCAGAGAACCACTGC	60
DB	1	AGTCCGCGTTTCGCGCGCTCGGTTGTCAACGGCGCGCCGCGCGCTCCATAGCGCAAC	60
QY	114	AGGACCGGAGTGGAGCGCGCGCAGCATGACGGCGCAGCGCCGCTCCATAGCGCAGTC	173
DB	61	AGGACCGGAGTGGAGCGCGCGCAGCATGACGGCGCAGCGCCGCTCCATAGCGCAGTC	120
QY	174	GGCAGCGTCCGCGCGGGGCGCGGGGGAGGAAATGCAACATGCGCAGCAATGGAAAC	233
DB	121	GGCAGCGTCCGCGCGGGGCGCGGGGGAGGAAATGCAACATGCGCAGCAATGGAAAC	180
QY	234	AGAACGCTGGCGTGTGGAGATATTTGAAACTGCGGAGCTGTGAGGAGAAATTTGAATCACA	293
DB	181	AGAACGCTGGCGTGTGGAGATATTTGAAACTGCGGAGCTGTGAGGAGAAATTTGAATCACA	240
QY	294	GGATCGGCTTAAATTTGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGCTACGCTTAAAA	353
DB	241	GGATCGGCTTAAATTTGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGCTACGCTTAAAA	300
QY	354	GCTGCTTGGCGATACCAGAGAAAATATCATGGCTACATGATGGCTAGGCGACCAATGATTT	413
DB	301	GCTGCTTGGCGATACCAGAGAAAATATCATGGCTACATGATGGCTAGGCGACCAATGATTT	360
QY	414	CATGTTTGTGAAAGAGGAGATGATCCAGATGGACTCATTCAGACAGAAATCTATTACCTTGC	473

ORIGIN	Query Match	Best Local Similarity	Score 526; DB 7; Length 674;	Matches 526; Conservative	Mismatches 0; Indels 0; Gaps 0;
Qy	1803	CCCTTTAGAGCATCACCTGTACGTAGTCAGTTTACCTGAATACCTCTGGAGAGGTGACAAGGCT	1862		
Db	1	CCCTTTAGAGCATCACCTGTACGTAGTCAGTTTACCTGAATACCTCTGGAGAGGTGACAAGGCT	60		
Qy	1863	GACTGACCGTGGGTACTTCACATTTCTTGTGCGATCAGTCAGCACTGTGACTTCTTTATAAG	1922		
Db	61	GACTGACCGTGGGTACTTCACATTTCTTGTGCGATCAGTCAGCACTGTGACTTCTTTATAAG	120		
Qy	1923	TAAGTATAGTAAACAGAGAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGA	1982		
Db	121	TAAGTATAGTAAACAGAGAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGA	180		
Qy	1983	AGATGACCCAATTTGCAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGAGATTCAGCAGTCC	2042		
Db	181	AGATGACCCAATTTGCAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGAGATTCAGCAGTCC	240		
Qy	2043	TCCTTCTGACTATATCTCCTCCAGAAATTTTCTCTTTTGAAGTACTACTCGGATTTACATT	2102		
Db	241	TCCTTCTGACTATATCTCCTCCAGAAATTTTCTCTTTTGAAGTACTACTCGGATTTACATT	300		
Qy	2103	GTATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTGCT	2162		
Db	301	GTATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTGCT	360		
Qy	2163	GTTTCATATATGTTGGTCTCAGGTGCGAGTTGGTGAATATTCGGTTTAAAGGATCAAGTA	2222		
Db	361	GTTTCATATATGTTGGTCTCAGGTGCGAGTTGGTGAATATTCGGTTTAAAGGATCAAGTA	420		
Qy	2223	TTTTCGGCTTGAATACCTTAGCTCTCTAGGTTATGTGTTGTAGTCATAGACAACAGGGG	2282		
Db	421	TTTTCGGCTTGAATACCTTAGCTCTCTAGGTTATGTGTTGTAGTCATAGACAACAGGGG	480		
Qy	2283	ATCCTGTGCAACGAGGGCTTAAATTGGAAGCGCCTTTAAATATAAA	2328		
Db	481	ATCCTGTGCAACGAGGGCTTAAATTGGAAGCGCCTTTAAATATAAA	526		

RESULT 86	565 bp	linear	EST 05-NOV-2005
DA694456	NT2NE2	Homo sapiens cDNA clone NT2NE2014882 5', mRNA	
LOCUS			
DEFINITION			
ACCESSION	DA694456		
VERSION	DA694456.1	GI:80772330	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 565)		
AUTHORS	Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaquri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kato, K., Yokoi, T., Kondoh, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.		
TITLE	Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes		
JOURNAL	Genome Res.	16 (1), 55-65	(2006)
PUBMED	16344560		
COMMENT	Contact: Takao Isogai		

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE 1 (bases 1 to 1717)
AUTHORS Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Teuritani, K., Takaguri, H., Ishii, S., Sugiyama, S., Saito, K., Isono, Y., Irie, R., Kuehida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isegai, T. and Sugano, S.
TITLE Diversification and Transcriptional Modulation: Large-scale Identification of Characterization of Putative Alternative Promoters of Human Genes
JOURNAL Promoters of Human Genes
PUBMED Genome Res. 16 (1), 55-65 (2006)
COMMENT 1634560
Contact: Takao Isegai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute (HRI); cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and HRI.

FEATURES
source Location/Qualifiers
1..717
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLAC1002059"
/tissue_type="placenta"
/clone_lib="PLAC1"
/note="Vector: pME18SFL3"

ORIGIN
Query Match 16.8%; Score 525; DB 9; Length 717;
Best Local Similarity 100.0%; Pred. No. 4,4e-272; Indels 0; Gaps 0;
Matches 525; Conservative 0; Mismatches 0;

QY 1482 TGAAGAACACACACATCTGGATTAATATCCATGACATCTTTCATGTTTCCCAAG 1541
DB 1 TGAAGAACACACACATCTGGATTAATATCCATGACATCTTTCATGTTTCCCAAG 60

QY 1542 TCACGAGAGGAATTCAGTTTATTTTTCCTCTGATTCGAAACAGGTTTCCTCATTT 1601
DB 61 TCACGAGAGGAATTCAGTTTATTTTTCCTCTGATTCGAAACAGGTTTTCCTCATTT 120

QY 1602 ATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGATCCAGTGGTGGCTGCC 1661
DB 121 ATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGATCCAGTGGTGGCTGCC 180

QY 1662 TGCTCCAGTGATTTCAAGTGTCCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATG 1721
DB 181 TGCTCCAGTGATTTCAAGTGTCCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATG 240

QY 1722 GGAAGTTCTTGGCGGCATGGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATA 1781
DB 241 GGAAGTTCTTGGCGGCATGGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATA 300

QY 1782 TTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTAAA 1841
DB 301 TTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTAAA 360

QY 1842 TCCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTACATTTCTTGGTGCATCACTCA 1901
DB 361 TCCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTACATTTCTTGGTGCATCACTCA 420

QY 1902 GCACCTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAGAATCCACACTGTGTCCCT 1961

DB 421 GCACCTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAGAATCCACACTGTGTCCCT 480
QY 1962 TTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACAAACAA 2006
DB 481 TTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACAAACAA 525

RESULT 88
DR002989
LOCUS
DEFINITION DR002989 Human fetal brain, large insert, pcMV expression library
TCl25354
Homo sapiens cDNA clone TCl25354, 5', similar to Homo sapiens dipeptidylpeptidase 8 (DPP8), transcript variant 2, mRNA sequence.
ACCESSION DR002989
VERSION DR002989.1 GI:66262862
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
REFERENCE 1 (bases 1 to 523)
AUTHORS Birkett, C., Cho, J., Gau, Y., Hamer, R., Kelly, S., Kovacs, K., Liu, L., Liu, X., Porter, J., Sachs, A., Shu, Y., Sun, Z., Wong, J., Wu, M., Zhang, X., Jay, G. and He, W.
TITLE High-throughput cloning of full-length human cDNAs directly from cDNA libraries optimized for large and rare transcripts
JOURNAL Unpublished (2005)
COMMENT Contact: Kovacs, KF
High Throughput cDNA Cloning
Origene Technologies, Inc. (www.origene.com)
6 Taft Court, Suite 100, Rockville, MD 20850, USA
Tel: 301 340 3188
Fax: 301 340 8606
Email: cdna@origene.com
This EST submission is part of an on-going human full-length cloning project at Origene Technologies, Inc.
Please contact Origene for access.
Origene Technologies, Inc.
6 Taft Ct, Suite 100
Rockville, MD 20850
Tel: (301) 340-3188
http://www.origene.com
Seq primer: pCMV6 5prime forward vector primer, Origene Technologies Inc.
Location/Qualifiers
1..523
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TCl25354"
/tissue_type="Fetal Brain"
/clone_lib="Human fetal brain, large insert, pcMV expression library"
/note="Organ: Fetal Brain; Vector: pCMV6-XL4; Site 1: Ecor1; Site 2: Xho1/SalI compatible end ligatio; Oligo-dt primed reverse transcription optimized for large and gc rich mRNA transcripts, cDNA size selection, optimized ligation for large inserts into mammalian expression vector, random clones selected for end sequence verification of full-length genes"

ORIGIN
Query Match 16.8%; Score 523; DB 9; Length 523;
Best Local Similarity 100.0%; Pred. No. 5.3e-271; Indels 0; Gaps 0;
Matches 523; Conservative 0; Mismatches 0;

QY 20 GCCAAGCGCGCTCTACTCCCGCTGCTTCTTAGTCCGCGTTCGCGCTGGTGT 79
DB 1 GCCAAGCGCGCTCTACTCCCGCTGCTTCTTAGTCCGCGTTCGCGCTGGTGT 60

QY 80 CACCGGCGCGCGCGCGGAGAACCACTGCAACACGAGACCGAGTGGAGCGCGCAGCA 139

QY 183 CGGGGGGGGGGGGGGAGAAATGCAACATGGCAGCAGCAATGGAACAGACAGCT 242
 Db 181 CGGGGGGGGGGGGGGAGAAATGCAACATGGCAGCAGCAATGGAACAGACAGCT 240
 QY 243 GGGTGTGAGATATTTGAAATCGGACTGTGAGGAGAAATTTGAATCAGAGATCGGCC 302
 Db 241 GGGTGTGAGATATTTGAAATCGGACTGTGAGGAGAAATTTGAATCAGAGATCGGCC 300
 QY 303 TAAATTCGAGCCCTTTTATTTGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGATGCTTGC 362
 Db 301 TAAATTCGAGCCCTTTTATTTGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGATGCTTGC 360
 QY 363 CGATACAGAAATATCATGGCTACATGATGCTTAAGCACCACATCATTCATGTTGT 422
 Db 361 CGATACAGAAATATCATGGCTACATGATGCTTAAGCACCACATCATTCATGTTGT 420
 QY 423 GAAGAGGAATGATTCAGATGGAACCTCATTCAGACAGAAATCTTATTCCTGCAATCTCG 482
 Db 421 GAAGAGGAATGATTCAGATGGAACCTCATTCAGACAGAAATCTTATTCCTGCAATCTCG 480
 QY 483 TGAGAACAGAGAAATACACTGTTTATTCGAAAT 519
 Db 481 TGAGAACAGAGAAATACACTGTTTATTCGAAAT 517

RESULT 94
 DA817784
 LOCUS
 DEFINITION DA817784 PEBLM2 Homo sapiens CDNA clone PEBLM2005870 5', mRNA EST 11-NOV-2005
 sequence.
 ACCESSION DA817784
 VERSION DA817784.1 GI:82170926
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 566)
 Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
 Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
 Ishii, S., Sugiyama, T., Saigo, K., Isono, Y., Irie, R., Kishida, N.,
 Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
 Tanabe, T., Nagai, K., Kikuchi, T., Takahashi, F., Fujii, A.,
 Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
 Diversification and Transcriptional Modulation of Large-scale
 Promoters of Human Genes
 Genome Res. 16 (1), 55-65 (2006)
 16344560
 Contact: Takao Isogai
 FLJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Katsurari, Kisearazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3925
 Fax: 81-438-52-3986
 Email: flj-cs@nifty.com
 NEDO human cDNA project (New Energy and Industrial Technology
 Developmental Organization, Japan); cDNA library construction;
 Helix Research Institute (HRI); 5'-end one pass sequencing; HRI,
 Research Association for Biotechnology (RAB) and Biotechnology
 Center, National Institute of Technology and Evaluation; 3'-end one
 pass sequencing: RAB.

FEATURES
 source
 1..566
 /location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PEBLM2005870"
 /cell_type="peripheral blood mononuclear cells
 (HPBMC5939)"
 /clone_lib="PEBLM2"
 /note="Vector: pME18SFL3; primary culture, peripheral

ORIGIN
 Query Match 16.5%; Score 515; DB 9; Length 566;
 Best Local Similarity 99.8%; Pred. No. 1.2e-266;
 Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 blood mononuclear cells"
 QY 19 GGCACAGGCGCTGCTACTGCGCGCTGCTCTTTAGTGCCTGCTTCCGCGCTGGGTTG 78
 Db 1 GGCACAGGCGCTGCTACTGCGCGCTGCTCTTTAGTGCCTGCTTCCGCGCTGGGTTG 60
 QY 79 TCACCCGCGCGCGCGCGCGAGGAGCACTGCAACAGAGCCGAGTGGAGCGCGCAGC 138
 Db 61 TCACCCGCGCGCGCGCGCGAGGAGCACTGCAACAGAGCCGAGTGGAGCGCGCAGC 120
 QY 139 ATGAAGCGCGCGAGCGCGCTCCATAGCGCAGCTCGGAGCGGTCCGCGCGCGCGGG 198
 Db 121 ATGAAGCGCGCGAGCGCGCTCCATAGCGCAGCTCGGAGCGGTCCGCGCGCGGGG 180
 QY 199 GAAGGAAATGCAACATGCGAGCAGCAATGGAACAGAAACAGCACTGGGTGTTGAGATATT 258
 Db 181 GAAGGAAATGCAACATGCGAGCAGCAATGGAACAGAAACAGCACTGGGTGTTGAGATATT 240
 QY 259 GAACTGCGGACTGTGAGGAGCAATTTGAATCAAGATCGGCTAAATTTGGAGCCTTTT 318
 Db 241 GAACTGCGGACTGTGAGGAGCAATTTGAATCAAGATCGGCTAAATTTGGAGCCTTTT 300
 QY 319 TATGTTGAGCGGTATTCCTGAGTCAAGCTTAAAGCTGCTTCCGATACAGGAAATAT 378
 Db 301 TATGTTGAGCGGTATTCCTGAGTCAAGCTTAAAGCTGCTTCCGATACAGGAAATAT 360
 QY 379 CATGCTACATGATGCTAAGGACCAACATGATTTCAATTTGTAAGAGGATGATCCA 438
 Db 361 CATGCTACATGATGCTAAGGACCAACATGATTTCAATTTGTAAGAGGATGATCCA 420
 QY 439 GATGAGCCTCATTCAGACAGAAATCTTATACCTTGCATGCTGTGAGAAACAGAGAAAT 498
 Db 421 GATGAGCCTCATTCAGACAGAAATCTTATACCTTGCATGCTGTGAGAAACAGAGAAAT 480
 QY 499 ACATGTTTATTTCTGAAATTTCCAAACTATCAATAGCAGCAGCTTAACTCTCTCT 558
 Db 481 ACATGTTTATTTCTGAAATTTCCAAACTATCAATAGCAGCAGCTTAACTCTCTCT 540
 QY 559 TGAAGCCTCTTTTGGATCTTTTCA 584
 Db 541 TGAAGCCTCTTTTGGATCTTTTCA 566

RESULT 95
 BE388695
 LOCUS
 DEFINITION 601286345F1 NIH_MGC_44 Homo sapiens CDNA clone IMAGE:3613139 5',
 mRNA sequence.
 ACCESSION BE388695
 VERSION BE388695.1 GI:9334060
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 658)
 NIH-MGC <http://mgs.nci.nih.gov/>;
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: the I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:

Dominidae; Homo.
 (bases 10 to 564)
 Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
 Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
 Shii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
 Monekawa, K., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
 Furukawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
 Murase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
 Diversification of Transcriptional Modulation: large-scale
 identification and Characterization of Putative Alternative
 Promoters of Human Genes
 Chromosome Res. 16 (1), 55-65 (2006)
 Contact: Takao Isogai
 FLJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 252-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: flj-cdna@nifty.com
 NEMO human cdna project (New Energy and Industrial Technology
 Developmental Organization, Japan); cdna library construction:
 Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
 Research Association for Biotechnology (RAB) and Biotechnology
 Center, National Institute of Technology and Evaluation; 3'-end one
 pass sequencing: RAB.
 Location/Qualifiers
 1. 564
 /organism="Homo sapiens"
 /mol_type="mrna"
 /db_xref="taxon:9606"
 /clone="NTONG2007406"
 /tissue_type="tongue"
 /clone_lib="NTONG2"
 /note="vector: PME18SFLJ3"

Pos	16.4%; 11 Similarity 563; Conservative	Score 513; Pred. No. 1.4e-265; 0; Mismatches	DB 9; Indels 1;	Length 564; Gaps 0;
1	AAGTGCTAAAGCCTCGAGGCGCAAGGCCGTGCTACTCTGCGCGCGTCTCTTAGTGCGCG	60		
1	AAGTGCTAAAGCCTCGAGGCGCAAGGCCGTGCTACTCTGCGCGCGTCTCTTAGTGCGCG	60		
61	CGTTGCGCGCCTGCGTCTGCAACCGCGCGCGCGCGAGGAGCACTTGCAAC CAGAGACCG	120		
61	CGTTGCGCGCCTGCGTCTGTCACCGCGCGCGCGCGAGGAGCACTTGCAAC CAGAGACCG	120		
21	GAGTGGAGCGCGCAGCATGAAGCGGCGCAGGCCCGCTCCATAGCGCAGCTCGGGACGG	180		
21	GAGTGGAGCGCGCAGCATGAAGCGGCGCAGGCCCGCTCCATAGCGCAGCTCGGGACGG	180		
81	TCCGGCGCGGGCGGGGAGGAAATGCAACATGCGCAGCAGCAATGGAACACAGAACAG	240		
81	TCCGGCGCGGGCGGGGAGGAAATGCAACATGCGCAGCAGCAATGGAACACAGAACAG	240		
441	CTGGGTGTTGAGATATTTGAAACTTCGGGACTGTGAGGAGAAATATTTGAATCACAGGATCGG	300		
441	CTGGGTGTTGAGATATTTGAAACTTCGGGACTGTGAGGAGAAATATTTGAATCACAGGATCGG	300		
301	CCTAAATTGGAGCGCTTTTATGTTGAGCGGTTATTCCTGGAGCTAGCTTAAAGACCTGCTT	360		
301	CCTAAATTGGAGCGCTTTTATGTTGAGCGGTTATTCCTGGAGCTAGCTTAAAGACCTGCTT	360		
361	GCCGATACCGAAAAATATCATGTGCTACATGATGGCTTAAGGCCACCATCATGATTTTCATGTTT	420		
361	GCCGATACCGAAAAATATCATGTGCTACATGATGGCTTAAGGCCACCATCATGATTTTCATGTTT	420		
361	GCCGATACCGAAAAATATCATGTGCTACATGATGGCTTAAGGCCACCATCATGATTTTCATGTTT	420		
421	GTGAAGAGGATGATCCAGATGCACTTCATTCAGACAGAAATCTATTACCTTGGCCATGTCT	480		
421	GTGAAGAGGATGATCCAGATGCACTTCATTCAGACAGAAATCTATTACCTTGGCCATGTCT	480		

```

http://image.llnl.gov
Plate: LICM273 row: i column: 12
High quality sequence stop: 609.
Location/Qualifiers
1. 658
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3613139"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NHGMC_44"
/note="Organ: uterus; Vector: pOTB7; Site: 1; XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN		Query Match	16.5%;	Score 515;	DB 7;	Length 658;
Best Local Similarity		100.0%;	Pred. No. 1.2e-266;	Mismatches 0;	Indels 0;	Gaps 0;
Matches 515;		Conservative				
QY	1811	AGCATCACCTGTACGTAGTACAGTACGTAACCTCTGGAGAGGTGACAAAGGCTGACTGACC	1870			
DB	71	AGCATCACCTGTACGTAGTACAGTACGTAACCTCTGGAGAGGTGACAAAGGCTGACTGACC	130			
QY	1871	GTGGCTACTCACATTTCTTGCTGCATCAGTCAGCACCTGTGACATTTCTTTATAAGTAAGTATA	1930			
DB	131	GTGGCTACTCACATTTCTTGCTGCATCAGTCAGCACCTGTGACATTTCTTTATAAGTAAGTATA	190			
QY	1931	GTAACCAAGAAGATCCACACTGTGTGTCCTTTACAAGCTATCAAGTCTCGAAGATGACC	1990			
DB	191	GTAACCAAGAAGATCCACACTGTGTGTCCTTTACAAGCTATCAAGTCTCGAAGATGACC	250			
QY	1991	CAACTTGCANAAACAAGAAATTTTGGGCCACCAATTTTGGATTTCAGACGTCCTCTTCTCTG	2050			
DB	251	CAACTTGCANAAACAAGAAATTTTGGGCCACCAATTTTGGATTTCAGACGTCCTCTTCTCTG	310			
QY	2051	ACTATACCTCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTACATTTGATGGGA	2110			
DB	311	ACTATACCTCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTACATTTGATGGGA	370			
QY	2111	TGCTCTACAGCCTCATGATCTACAGCCTCGAAAGAATATCCCTACTGTGCTGTTCAATAT	2170			
DB	371	TGCTCTACAGCCTCATGATCTACAGCCTCGAAAGAATATCCCTACTGTGCTGTTCAATAT	430			
QY	2171	ATGGTGGTCTCCAGTGCAGTTGGTGAATAATTCGGTTTAAAGGAGTCAGTATTCGGCT	2230			
DB	431	ATGGTGGTCTCCAGTGCAGTTGGTGAATAATTCGGTTTAAAGGAGTCAGTATTCGGCT	490			
QY	2231	TGAATACCTACCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCCTGTCT	2290			
DB	491	TGAATACCTACCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCCTGTCT	550			
QY	2291	ACCGGGGCTTAAATTTGAAGCGCGCTTTTAAATAT	2325			
DB	551	ACCGGGGCTTAAATTTGAAGCGCGCTTTTAAATAT	585			

RESULT	96
LOCUS	DA766261
DEFINITION	DA766261 NTONG2 Homo sapiens cDNA clone NTONG2007406 5', mRNA linear EST 09-NOV-2005 sequence.
ACCESSION	DA766261
VERSION	DA766261.1 GI:81279987
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

QY 481 GGTGAGAACAGAGAAATACACCTGTTTATCTGAAATTCCTCAATAGAGCA 540
 Db 481 GGTGAGAACAGAGAAATACACCTGTTTATCTGAAATTCCTCAATAGAGCA 540
 QY 541 GCAGCTCTTAATGCTCTCTTGGGAAG 564
 Db 541 GCAGCTCTTAATGCTCTCTTGGGAAG 564

RESULT 97
 DA033029
 LOCUS
 DEFINITION DA033029 ASTRO2 Homo sapiens cDNA clone ASTRO2017889 5', mRNA
 sequence.
 ACCESSION DA033029
 VERSION
 KEYWORDS
 SOURCE EST.
 ORGANISM Homo sapiens (human)

REFERENCE
 AUTHORS Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saiko, K., Isono, Y., Irie, R., Kishida, N., Yoneyama, T., Ohtsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
 TITLE Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
 JOURNAL Genome Res. 16 (1), 55-65 (2006)
 PUBMED 16344560
 COMMENT Contact: Takao Isogai
 FLJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: flj-cdna@nifty.com
 NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

FEATURES
 source
 Location/Qualifiers
 1..563
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="ASTRO2017889"
 /cell_type="normal astrocytes (NHAS732)"
 /clone_lib="ASTRO2"
 /note="Vector: pME18SFL3"

ORIGIN
 Query Match 16.4%; Score 511; DB 9; Length 563;
 Best Local Similarity 99.8%; Pred No 1.7e-264;
 Matches 561; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 AAAGCTCCGAGGCCAGGCGCTGCTACTGCGCGCTGCTTCTTAGTGGCGCTTCGC 67
 Db 2 AAAGCTCCGAGGCCAGGCGCTGCTACTGCGCGCTGCTTCTTAGTGGCGCTTCGC 61
 QY 68 GCCTCTGGGTCTACCGCGCGCGCGCGAGGAGCCATGCAACAGGACCGGAGTGA 127
 Db 62 GCCTCTGGGTCTACCGCGCGCGCGCGAGGAGCCACTGCAACAGGACCGGAGTGA 121
 QY 128 GCGCGCGCAGCATGAAGCGCGAGCGCGCTCCATAGCGACGTCGGAGCGTCCGGGC 187
 |||||

Db 122 GCGCGCGCAGCATGAAGCGCGCGCGCTCCATAGCGACGTCGGAGCGTCCGGGC 181
 QY 188 GCGCGCGCGCGAGGAGAAATGCAACATGCGAGCGCAATGGAAACAGAACAGCTGGGTG 247
 Db 182 GCGCGCGCGCGAGGAGAAATGCAACATGCGAGCGCAATGGAAACAGAACAGCTGGGTG 241
 QY 248 TTGAGATATTTGAAACTGCGGACTGTGGAGGAATATTGAATCAAGGATCGGCTAAAT 307
 Db 242 TTGAGATATTTGAAACTGCGGACTGTGGAGGAATATTGAATCAAGGATCGGCTAAAT 301
 QY 308 TGGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAAAGCTGCTCCCGATA 367
 Db 302 TGGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAAAGCTGCTCCCGATA 361
 QY 368 CAGAAAATATCATGCTACATGATGGCTAAGGACCAACATGATTTTCATGTTGTGAAGA 427
 Db 362 CAGAAAATATCATGCTACATGATGGCTAAGGACCAACATGATTTTCATGTTGTGAAGA 421
 QY 428 GGATGATCCAGATGGACCTCAATTCAGACAGAATCTATTACCTTGCCATCTCTGGTGAGA 487
 Db 422 GGATGATCCAGATGGACCTCAATTCAGACAGAATCTATTACCTTGCCATCTCTGGTGAGA 481
 QY 488 ACAGAGAAATACACTGTTTATTCTGAAATTCCTCAAACTATCAATAGGACGAGTCT 547
 Db 482 ACAGAGAAATACACTGTTTATTCTGAAATTCCTCAAACTATCAATAGGACGAGTCT 541
 QY 548 TAATGCTCTCTTGGAGCGCTCT 569
 Db 542 TAATGCTCTCTTGGAGCGCTCT 563

RESULT 98
 DB020491
 LOCUS
 DEFINITION DB020491 TEST12 Homo sapiens cDNA clone TEST12000701 5', mRNA
 sequence.
 ACCESSION DB020491
 VERSION
 KEYWORDS
 SOURCE EST.
 ORGANISM Homo sapiens (human)

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 559)
 Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saiko, K., Isono, Y., Irie, R., Kishida, N., Yoneyama, T., Ohtsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
 TITLE Diversification and Characterization of Putative Alternative Promoters of Human Genes
 JOURNAL Genome Res. 16 (1), 55-65 (2006)
 PUBMED 16344560
 COMMENT Contact: Takao Isogai
 FLJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: flj-cdna@nifty.com
 NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

FEATURES
 source
 Location/Qualifiers
 1..559
 /organism="Homo sapiens"
 /mol_type="mRNA"

RESULT 99	DB242468	575 bp	mRNA	linear	EST 04-DEC-2005
LOCUS	DB242468				
DEFINITION	TSTOM2 Homo sapiens cDNA clone TSTOM2000881 5', mRNA				
ACCESSION	DB242468				
VERSION	DB242468				
KEYWORDS	EST				
SOURCE	DB242468.1	GI:83219482			
ORGANISM	Homo sapiens (human)				
REFERENCE	1 (bases 1 to 575)				
AUTHORS	Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamaashita, R., Yamamoto, J., Sekine, M., Tsuricani, K., Wakisugi, H., Iehi, S., Sugiyama, T., Saito, K., Isono, Y., Irie, K., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Iehida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.				
TITLE	Diversification of Transcriptional Modulation: Large-scale Identification of Human Genes Promoters of Human Genes				
JOURNAL	Genome Res. 16 (1), 55-65 (2006)				
PUBLISHED	16344560				

sequence.	377	ATCATGGCTACATGATGCTTAAGGCAACACATGATTTTCATGTTTGTGAAGAGGAATGATC	436
DA523081	Db		
EST	361	ATCATGGCTACATGATGCTTAGGCAACACATGTTTCATGTTTGTGAAGAGGAATGATC	420
Hom sapiens	Qy		
Hom sapiens (human)	Db		
Eukaryota	437	CAGATGGACCTCATTTCAGACAGATCTTATTACCTTCCCATGCTTGTGTGAGAACAGAGAAA	496
Mammalia	421		
Hominiidae; Homo	Qy		
1 (bases 1 to 560)	Db		
Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saiko,K., Isono,Y., Irie,R., Kuehida,N., Yoneyama,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S. Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes	497	ATACACTGTTTTTATTCGAAATCCCAAACTACTCAATAGACAGCAGTCTTAATGCTCT	556
Genome Res. 16 (1), 55-65 (2006)	Db		
Contact: Takao Isogai	557	CTTGGAAAGCCTCTTTTGGAT	576
FLJ Project (HRI Team)	541	CTTGGAAAGCCTCTTTTGGAT	560
Helix Research Institute	RESULT 101		
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan	DA437193		
Tel: 81-438-52-3975	LOCUS		
Fax: 81-438-52-3986	DEFINITION		
Email: flj-cdna@nifty.com	DA437193		
NEO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.	ACCESSION		
Location/Qualifiers	VERSION		
1. 560	DA437193.1		
/organism="Homo sapiens"	KEYWORDS		
/mol_type="mRNA"	SOURCE		
/db_xref="taxon:9606"	ORGANISM		
/clone="FEBA2013746"			
/tissue_type="brain"			
/dev_stage="fetal"			
/clone_lib="FEBA2"			
/notes="Vector: pME18SFL3"			
	REFERENCE		
	AUTHORS		
	1 (bases 1 to 611)		
	Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saiko,K., Isono,Y., Irie,R., Kuehida,N., Yoneyama,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S. Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes		
	Genome Res. 16 (1), 55-65 (2006)		
	Contact: Takao Isogai		
	FLJ Project (HRI Team)		
	Helix Research Institute		
	2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan		
	Tel: 81-438-52-3975		
	Fax: 81-438-52-3986		
	Email: flj-cdna@nifty.com		
	NEO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.		
	Location/Qualifiers		
	1. 560		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="FEBA2013746"		
	/tissue_type="brain"		
	/dev_stage="fetal"		
	/clone_lib="FEBA2"		
	/notes="Vector: pME18SFL3"		
	FEATURES		
	source		
	16.3%; Score 509; DB 9; Length 560;		
	Best Local Similarity 99.8%; Pred. No. 2.1e-263;		
	Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
	17 GAGGCAAGCCCGCTGCTACTACCGCCGCTGCTTCTTAGTCGCCGCTTCGCCGCTGGGT 76		
	1 GAGGCCAAGCCCGCTGCTACTACTCGCCGCTGCTTCTTAGTCGCCGCTTCGCCGCTGGGT 60		
	77 TGTACACCGCGCGCGCGAGGAAGCCACTGCAACACAGCAGCGATGAGCGCGCGCA 136		
	61 TGTACACCGCGCGCGCGAGGAAGCCACTGCAACACAGCAGCGATGAGCGCGCGCA 120		
	137 GCATGAAGCGCGCGCGCGCGCTTCATAGCGCAGCTCGGACGCTCGGCGCGCGCGG 196		
	121 GCATGAAGCGCGCGCGCGCGCTTCATAGCGCAGCTCGGACGCTCGGCGCGCGCGG 180		
	197 GCGAAGGAAATGCAACATGCGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATAT 256		
	181 GCGAAGGAAATGCAACATGCGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATAT 240		
	257 TTGAACCTGCGGATCTGTGAGGAGAAATTCGAATCAGAGTTCGCCCTAAATTTGAGCCTT 316		
	241 TTGAACCTGCGGATCTGTGAGGAGAAATTTGAATCAGAGTTCGCCCTAAATTTGAGCCTT 300		
	317 TTTATGTTGAGCGGTATTTCTCGAGTACGTTTAAAGCTGTTTCCCGATACCAAGAAAT 376		
	301 TTTATGTTGAGCGGTATTTCTCGAGTACGTTTAAAGCTGTTTCCCGATACCAAGAAAT 360		
	ORIGIN		
	Query Match		
	Best Local Similarity 99.8%; Score 509; DB 9; Length 560;		
	Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
	17 GAGGCAAGCCCGCTGCTACTACCGCCGCTGCTTCTTAGTCGCCGCTTCGCCGCTGGGT 76		
	1 GAGGCCAAGCCCGCTGCTACTACTCGCCGCTGCTTCTTAGTCGCCGCTTCGCCGCTGGGT 60		
	77 TGTACACCGCGCGCGCGAGGAAGCCACTGCAACACAGCAGCGATGAGCGCGCGCA 136		
	61 TGTACACCGCGCGCGCGAGGAAGCCACTGCAACACAGCAGCGATGAGCGCGCGCA 120		
	137 GCATGAAGCGCGCGCGCGCGCTTCATAGCGCAGCTCGGACGCTCGGCGCGCGCGG 196		
	121 GCATGAAGCGCGCGCGCGCGCTTCATAGCGCAGCTCGGACGCTCGGCGCGCGCGG 180		
	197 GCGAAGGAAATGCAACATGCGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATAT 256		
	181 GCGAAGGAAATGCAACATGCGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATAT 240		
	257 TTGAACCTGCGGATCTGTGAGGAGAAATTCGAATCAGAGTTCGCCCTAAATTTGAGCCTT 316		
	241 TTGAACCTGCGGATCTGTGAGGAGAAATTTGAATCAGAGTTCGCCCTAAATTTGAGCCTT 300		
	317 TTTATGTTGAGCGGTATTTCTCGAGTACGTTTAAAGCTGTTTCCCGATACCAAGAAAT 376		
	301 TTTATGTTGAGCGGTATTTCTCGAGTACGTTTAAAGCTGTTTCCCGATACCAAGAAAT 360		
	ORIGIN		
	Query Match		
	Best Local Similarity 9		

Db	1	AGTGGCGGTTTGGCGGCTTGGGTTGTACCGGCGCGCGCGGAGGAAGCCAATGCAACC	60
Qy	114	AGGACCGGAGTGGAGCGCGCAGCATGAAGGGCGGAGCGCCGCTCCATAGCGCAGTC	173
Db	61	AGGACCGGAGTGGAGCGCGCAGCATGAAGCGCGCAGCGCGCTCCATAGCGCGCTC	120
Qy	174	GGGACGGTCCGGCGGCGCGGGGGGAGGAAATGCAACATGTCAGCAGCAATGGAAAC	233
Db	121	GGGACGGTCCGGCGGCGCGGGGGAGGAAATGCAACATGTCAGCAGCAATGGAAAC	180
Qy	234	AGAACGAGCTGGGTGTTGAGATATTTGAAACTGCGGAGCTGTGAGGAGAAATTTGAATCACA	293
Db	181	AGAACGAGCTGGGTGTTGAGATATTTGAAACTGCGGAGCTGTGAGGAGAAATTTGAATCACA	240
Qy	294	GGATCGGCTTAATTCGAGCGCTTTTATGCTTCAGCGGTATTCCTGGAGTCAGCTTAAAAA	353
Db	241	GGATCGGCTTAATTCGAGCGCTTTTATGCTTCAGCGGTATTCCTGGAGTCAGCTTAAAAA	300
Qy	354	GCTGCTTCGCGATACAGAAAAATATCATGGCTACATGATGGCTANGCGACCAACATGATTT	413
Db	301	GCTGCTTCGCGATACAGAAAAATATCATGGCTACATGATGGCTANGCGACCAACATGATTT	360
Qy	414	CATGTTTGTGAAGGNAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGC	473
Db	361	CATGTTTGTGAAGGNAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGC	420
Qy	474	CATGCTCTGGTGAGAACAGAGAAATACATGTTTTTATCTGAAATCCCAAACTATCAA	533
Db	421	CATGCTCTGGTGAGAACAGAGAAATACATGTTTTTATCTGAAATCCCAAACTATCAA	480
Qy	534	TAGAGCAGCAGTCTTAATGCTCTCTGGAGCGCTCTTTTGGATCTTTTTTCAGGCAACACT	593
Db	481	TAGAGCAGCAGTCTTAATGCTCTCTGGAGCGCTCTTTTGGATCTTTTTTCAGGCAACACT	540
Qy	594	GGACTATCGAATGTATTTCTCGAGAAGAGAACTATTAAAGAAAGAAAAACGCAATTGGAC	653
Db	541	GGACTATCGAATGTATTTCTCGAGAAGAGAACTATTAAAGAAAGAAAAACGCAATTGGAC	600
Qy	654	AGTCGGAATTG	664
Db	601	AGTCGGAATTG	611

RESULT 102	DB041170	547 bp	linear	EST 08-NOV-2005
LOCUS	DB041170		mRNA	
DEFINITION	DB041170	TSST22 Homo sapiens cDNA clone	TSST22028180 5', mRNA	
ACCESSION	DB041170			
VERSION	DB041170	1	GI:81223330	
KEYWORDS	EST.			
SOURCE	Homo sapiens	(human)		
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Euarchontoglires; Primates; Catearrhini;			
	Hominidae; Homo.			
REFERENCE	1	(Bases 1 to 547)		
AUTHORS	Kimura K., Wakamatsu A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ighii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatauma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A. and Sugano, S., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, F. and Sugano, S.			
TITLE	Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes			
JOURNAL	Genome Res.	16 (1), 55-65	(2006)	
PUBMED	16344560			
COMMENT	Contact: Takao Isogai FLJ Project (HRI Team) Helix Research Institute 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan Tel.: 81-438-52-3975			

Fax: 81-438-52-3986
Email: fij-cdn@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology one
Center, National Institute of Technology and Evaluation; 3'-end
pass sequencing: RAB.

Location/Qualifiers
1. .547
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TEST12028180"
/tissue_type="testis"
/clone_lib="TEST12"
/note="Vector: pME18SFL3"

FEATURES
source

ORIGIN	Query Match Best Local Similarity Matches	16.3%; 100.0%; 508;	Score 508; Pred. No. 7.2e-263; Conservative 0;	DB 9; Indels 0; Mismatches 0;	Length 547;
Qy	1332	TGCTTGCGTCCATCTACTAGATCGCTCCAGACATCGCTACAGATAGTGTGATCTCACC	1391		
Db	40	TGCTTGCGTCCATCTACTAGATCGCTCCAGACATCGCTACAGATAGTGTGATCTCACC	99		
Qy	1392	TGAATTATTATTTATCCCAAGTAAAGATGATGTTATGGAAAGGCAGAGACATCAATTGAGTCAGT	1451		
Db	100	TGAATTATTATTTATCCCAAGTAAAGATGATGTTATGGAAAGGCAGAGACATCAATTGAGTCAGT	159		
Qy	1452	GCCTGATTCGTGACGCCCACTAAATTATCTATGAAGAAAACACAGACATCTCGGATAAATAT	1511		
Db	160	GCCTGATTCGTGACGCCCACTAAATTATCTATGAAGAAAACACAGACATCTCGGATAAATAT	219		
Qy	1512	CCATGACATCTTTTCATGTTTTTCCCAAAAGTCACGAAGGAGAAATGAGTATTTATTTTGC	1571		
Db	220	CCATGACATCTTTTCATGTTTTTCCCAAAAGTCACGAAGGAGAAATGAGTATTTATTTTGC	279		
Qy	1572	CTCTGAATGCAAAACAGGTTTCGGTCATTTATACAAATTAACATCTATTTTAAAGGAAAG	1631		
Db	280	CTCTGAATGCAAAACAGGTTTCGGTCATTTATACAAATTAACATCTATTTTAAAGGAAAG	339		
Qy	1632	CAAAATATAACGATCCAGTGGTGGCTGCCCTGCCAAGTGATTTCAAGTGTCCTATCAA	1691		
Db	340	CAAAATATAACGATCCAGTGGTGGCTGCCCTGCCAAGTGATTTCAAGTGTCCTATCAA	399		
Qy	1692	AGAGGAGATGCAATTAACCAAGTGTGTAATGGGAAGTTCTTGGCCGGCATGGATCTAATAT	1751		
Db	400	AGAGGAGATGCAATTAACCAAGTGTGTAATGGGAAGTTCTTGGCCGGCATGGATCTAATAT	459		
Qy	1752	CCAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGGAGGCCACCAAGACATCCCTTTAGA	1811		
Db	460	CCAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGGAGGCCACCAAGACATCCCTTTAGA	519		
Qy	1812	GCATCACCTGTACGTAGTCAGTTACGTA	1839		
Db	520	GCATCACCTGTACGTAGTCAGTTACGTA	547		

RESULT	103
DA370407	
LOCUS	559 bp mRNA linear EST 06-NOV-2005
DEFINITION	DA370407 BRTHA1 Homo sapiens cDNA clone BRTHAL000306 5', mRNA sequence.
ACCESSION	DA370407
VERSION	DA370407.1 GI:80866355
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
	Hominidae; Homo.


```
QY 1634 ATATATAACGATCCAGTGGTGGCTGCTCCAGTGATTTCAAGTGTCTCTATCAAG 1693
Db |||||||
QY 1694 AGGAGATAGCAATACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCC 1753
Db |||||||
QY 1754 AAGTGTAGTGAAGTCCAGAGGCTGGTATATTTTGAAGCCACCAAGACTCCCTTTAGAGC 1813
Db |||||||
QY 301 AAGTTTGAAGTGAAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGC 360
QY 1814 ATCACTCTAGTGTAGTGTACGTAAATCCTGGAGAGGTGAACAAGCTGACTGACCCGTG 1873
Db |||||||
QY 361 ATCACTCTAGTGTAGTGTACGTAAATCCTGGAGAGGTGACAAGGCTGACTGACCCGTG 420
QY 1874 GCTACTCATATCTTGTGTCATGATCAGTCAGCACTGTGACTTCTTTAAGTAAGTATAGTA 1933
Db |||||||
QY 421 GCTACTCATATCTTGTGTCATGATCAGTCAGCACTGTGACTTCTTTAAGTAAGTATAGTA 480
QY 1934 ACCAGAAGATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAA 1993
Db |||||||
QY 481 ACCAGAAGATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAA 540
QY 1994 CTTGCAAAACAAAGGAATT 2012
Db |||||||
541 CTTGCAAAACAAAGGAATT 559
```

RESULT 105

```
DA450229 LOCUS DA450229 CTONG2 Homo sapiens cDNA clone CTONG2022308 5', mRNA EST 06-NOV-2005
DEFINITION sequence.
```

ACCESSION DA450229

VERSION DA450229.1 GI:80908670

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamaehita, R., Yamamoto, J., Sekine, M., Teuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S. Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes

Genome Res. 16 (1), 55-65 (2006)

16344560

Contact: Takao Isogai

FLJ Project (HRI Team)

Helix Research Institute

2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: flj-cdna@nifty.com

NEO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

FEATURES

source

1..553

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CTONG2022308"

/tissue_type="tongue, tumor tissue"
/clone_lib="CTONG2"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 16.1%; Score 502; DB 9; Length 553;
Best Local Similarity 99.8%; Pred. No. 1.3e-259;
Matches 552; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAGTGTAAAGCCTTCGAGGCCAAGCCGCTGCTACTGCCCGCCGCTTCTTATGTCGG 60
Db |||||||

QY 1 AAGTGTAAAGCCTTCGAGGCCAAGCCGCTGCTACTGCCCGCCGCTTCTTATGTCGG 60
Db |||||||

QY 61 CGTTCCGCCCTTGGGTTGTACCCGGCGCGCGCGGAGGAAGCACTGCAACACGAGACCG 120
Db |||||||

QY 61 CGTTCCGCCCTTGGGTTGTACCCGGCGCGCGCGGAGGAAGCACTGCAACACGAGACCG 120
Db |||||||

QY 121 GAGTGAGCGCGCGCAGCATGAAAGCGCGCGCGCTCCATAGCGCACGTCGGGACGG 180
Db |||||||

QY 121 GAGTGAGCGCGCGCAGCATGAAAGCGCGCGCGCTCCATAGCGCACGTCGGGACGG 180
Db |||||||

QY 181 TCCGGCGCGCGCGCGGAGGAAATGCAACATGGCAGCAGCAATGGAACACGAGACAG 240
Db |||||||

QY 181 TCCGGCGCGCGCGCGGAGGAAATGCAACATGGCAGCAGCAATGGAACACGAGACAG 240
Db |||||||

QY 241 CTGGGTGTTGAGATATTTGAAACTCGCGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300
Db |||||||

QY 241 CTGGGTGTTGAGATATTTGAAACTCGCGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300
Db |||||||

QY 301 CCTAAATTCGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTT 360
Db |||||||

QY 301 CCTAAATTCGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTT 360
Db |||||||

QY 361 GCCGATACCAGAAAATATCATGGCTACATGATGGCTAAGGCACACATGATGTTT 420
Db |||||||

QY 361 GCCGATACCAGAAAATATCATGGCTACATGATGGCTAAGGCACACATGATGTTT 420
Db |||||||

QY 421 GTGAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTATACCTTGCCATGCT 480
Db |||||||

QY 421 GTGAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTATACCTTGCCATGCT 480
Db |||||||

QY 481 GGTGAGAACAGAGAAAATACACTGTTTTATTTCTGAAATTCCTCAAACTATCAATAGAC 540
Db |||||||

QY 481 GGTGAGAACAGAGAAAATACACTGTTTTATTTCTGAAATTCCTCAAACTATCAATAGAC 540
Db |||||||

QY 541 GCAGTCTTAATGC 553
Db |||||||

QY 541 GCAGTCTTAATGC 553
Db |||||||

RESULT 106

DA831627

LOCUS

DEFINITION

sequence.

DA831627

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 631)

REFERENCE

AUTHORS

Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,

Yamaehita, R., Yamamoto, J., Sekine, M., Teuritani, K., Wakaguri, H.,

Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N.,

Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,

Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,

Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.

Diversification of Transcriptional Modulation: Large-scale

Identification and Characterization of Putative Alternative

Promoters of Human Genes

TITLE

JOURNAL
PUBMED
COMMENT

Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
HRI human cdna project; 5'- & 3'-end one pass sequencing: Helix
Research Institute (HRI); cDNA library construction: department of
Virology, Institute of Medical Science, University of Tokyo, and
HRI.

Location/Qualifiers
1. .631
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLAC1007416"
/tissue_type="placenta"
/clone_lib="PLACE1"
/notes="Vector: pME18SFL3"

Query Match
Best Local Similarity 99.8%; Pred. No. 4.6e-259; Length 631;
Matches 551; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 832 GCTGATCCAGACTGGATCGTTTATACATAGCAACGATATTGGATATCTAACATCGTA 891
Db
QY 892 ACCAGAGAGAGAGAGACTCTATATGTGCACAACTAGCTAGCCAACTGGAAGAGAT 951
Db 97 ACCAGAGAGAGAGAGACTCTATATGTGCACAACTAGCTAGCCAACTGGAAGAGAT 156
QY 952 GCAGATCAGCTGGAGTGCCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTCTGC 1011
Db 157 GCAGATCAGCTGGAGTGCCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTCTGC 216
QY 1012 TATTGTTGTTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTAGATTCTTA 1071
Db 217 TATTGTTGTTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTAGATTCTTA 276
QY 1072 TATGAAGAAATGATGATCTGAGGTGGAATTAATTCATGTTACATCCCTATGTTGAA 1131
Db 277 TATGAAGAAATGATGATCTGAGGTGGAATTAATTCATGTTACATCCCTATGTTGAA 336
QY 1132 ACAGGAGGGCAGATTCAITCCGTTATCTTAAACAGGTACAGCAATCTTAAGTCACT 1191
Db 337 ACAGGAGGGCAGATTCAITCCGTTATCTTAAACAGGTACAGCAATCTTAAGTCACT 396
QY 1192 TTTAAGATGTCAGAAATTAATGATGCTGAGGAGGATCATAGATCTATAGATAAG 1251
Db 397 TTTAAGATGTCAGAAATTAATGATGCTGAGGAGGATCATAGATCTATAGATAAG 456
QY 1252 GAATTAATCAACCTTTGAGATTTCTATTGAGGAGTGTGATATATGCGAGAGCTGGA 1311
Db 457 GAATTAATCAACCTTTGAGATTTCTATTGAGGAGTGTGATATATGCGAGAGCTGGA 516
QY 1312 TGGACTCCTGAGGAGAAATATGCTGGTCCATCTCTAGATCGCTCCAGATCTGCGCTTA 1371
Db 517 TGGACTCCTGAGGAGAAATATGCTGGTCCATCTCTAGATCGCTCCAGATCTGCGCTTA 576
QY 1372 CAGATAGTGTG 1383
Db 577 CAGATAGTGTG 588

RESULT 107
BI546487
LOCUS
DEFINITION 603188793F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5260151 5',
EST 05-SBP-2001

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BI546487
BI546487.1 GI:15433799
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo
1 (bases) (CO 542)
NIH-MGC http://mhc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Straubeberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue: Placenta; Recombinant: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Tohyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LMAM1655 row: j column: 24
High quality sequence stop: 640.

FEATURES
source

1. .642
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5260151"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_95"
/notes="Organ: brain; Vector: pBluescriptR (modified
pUC19); Site 1: BamHI; Site 2: SalI-XhoI
5'-ATTTTATTTTATTTTATTTT-3' site-selected for average
insert size 2.5 kb and normalized to 10^5. This is a
primary library enriched for full-length clones and
constructed using the CapTrapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this
is a NIH_MGC Library."

ORIGIN

Query Match
Best Local Similarity 100.0%; Pred No. 1.6e-258; Length 642;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 AAAGCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCTGCTTCTTAGTGGCGCTTCGC 67
Db 13 AAAGCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCTGCTTCTTAGTGGCGCTTCGC 72
QY 68 CGCTGCGTGTGTCACCG 127
Db 73 CGCTGCGTGTGTCACCG 132
QY 128 GCGCGCGCAGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 187
Db 133 GCGCGCGCAGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 192
QY 188 GCG 247
Db 193 GCG 252
QY 248 TTGAGATATTTGAAATCTCGGACTCTGAGGAGATATTGAATCACAGATCGGCTTAAT 307
Db 253 TTGAGATATTTGAAATCTCGGACTCTGAGGAGATATTGAATCACAGATCGGCTTAAT 312
QY 308 TGAGAGCTTTTATGTTGAGCGGTATTCCTGAGGTCAAGTCTTAAAGCTCTTGGCGATA 367
Db 313 TGAGAGCTTTTATGTTGAGCGGTATTCCTGAGGTCAAGTCTTAAAGCTCTTGGCGATA 372

```

183 AAATGCAACATGGCGAGCAATGGAACAGACAGCTGGTGTGAGATATTTGAAACT 242
265 GCGGACTGTGAGGAGAAATTAATCAAGATCGGCTAAATTTGGAGCCTTTTATGTT 324
243 GCGGACTGTGAGGAGAAATTAATCAAGATCGGCTAAATTTGGAGCCTTTTATGTT 302
325 GAGCGTATTCCTGAGTCAAGTAAAGAGCTGTCGCCATACACAGAAAATATCATGGC 384
303 GAGCGTATTCCTGAGTCAAGTAAAGAGCTGTCGCCATACACAGAAAATATCATGGC 362
385 TACATGATGCTTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGA 444
363 TACATGATGCTTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGA 422
445 CCTCATTCAGACAGAAATCTTATACCTTCCATGCTGCTGAGAGAAACATACACTG 504
423 CCTCATTCAGACAGAAATCTTATACCTTCCATGCTGCTGAGAGAAACATACACTG 482
505 TTTTATTCCTGAAATTC 520
483 TTTTATTCCTGAAATTC 498

RESULT 109
AA496257/c
LOCUS
DEFINITION
  aa24b08.s1 NCI CGAP GCBI Homo sapiens cDNA clone IMAGE:814167 3',
  similar to SW:DPPE_HUMAN P42658 DIPEPTIDYL PEPTIDASE IV LIKE
  PROTEIN ;, mRNA sequence.
ACCESSION
  AA496257
VERSION
  AA496257.1 GI:2229578
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Hominidae; Homo.
REFERENCE
  1 (bases 1 to 500)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgaps-remail.nih.gov
  Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
  Ph.D., Gerald Marti, M.D.
  CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
  Bonaldo, Ph.D.
  CDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html

FEATURES
  source
  1..500
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="GDB:6032116"
  /db_xref="taxon:9606"
  /clone="IMAGE:814167"
  /tissue_type="germinal center B cell"
  /lab_host="DH10B"
  /notes="Vector: pT73D-PacI; Site 1: Not I; Site 2: Eco RI;
  1st strand cDNA was prepared from human tonsillar cells
  enriched for germinal center B cells by flow sorting
  (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr.
  David Allman (NCI) and Dr. Gerald Marti (CBER). CDNA
  synthesis was primed with a Not I - oligo(dT) primer"

Possible reversed clone: similarity on wrong strand
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 405.
Location/Qualifiers
  1..500
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="GDB:6032116"
  /db_xref="taxon:9606"
  /clone="IMAGE:814167"
  /tissue_type="germinal center B cell"
  /lab_host="DH10B"
  /notes="Vector: pT73D-PacI; Site 1: Not I; Site 2: Eco RI;
  1st strand cDNA was prepared from human tonsillar cells
  enriched for germinal center B cells by flow sorting
  (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr.
  David Allman (NCI) and Dr. Gerald Marti (CBER). CDNA
  synthesis was primed with a Not I - oligo(dT) primer"

```

```

368 CCAGAAAATATCATGGCTACATGATGCTAAGGCACACATGATTTTCATGTTTGAAGA 427
373 CCAGAAAATATCATGGCTACATGATGCTAAGGCACACATGATTTTCATGTTTGAAGA 432
428 GGAATGATCCAGTGAAGTCAATTCAGACAGAAATCTTATACCTTGCATGCTGCTGAGGA 487
433 GGAATGATCCAGTGAAGTCAATTCAGACAGAAATCTTATACCTTGCATGCTGCTGAGGA 492
488 ACAGAGAAAATACACTGTTT 507
493 ACAGAGAAAATACACTGTTT 512

CN427226
17000455364785 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
CN427226
CN427226.1 GI:47414820
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
  1 (bases 1 to 704)
  Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
  Lebkowski, J. and Stanton, L.W.
  Transcriptome characterization elucidates signaling networks that
  control human ES cell growth and differentiation
  Nat. Biotechnol. 22 (6), 707-716 (2004)
  15146197
  Contact: Brandenberger R
  Regenerative Medicine
  Genon Corporation
  230 Constitution Drive, Menlo Park, CA 94025, USA
  Tel: 650 473 8658
  Fax: 650 473 7760
  Email: rbrandenberger@genon.com
  Insert Length: 704 Std Error: 0.00.
  Location/Qualifiers
  1..704
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /tissue_type="embryonic stem cells, cell lines H1, H7, and
  H9"
  /clone_lib="GEN_ES"
  /note="Oligo dT primed, full-length enriched cDNA library
  from undifferentiated hES cell lines H1 (p32), H7 (p29),
  and H9 (p26) maintained in feeder-free conditions"

FEATURES
  source
  1..704
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /tissue_type="embryonic stem cells, cell lines H1, H7, and
  H9"
  /clone_lib="GEN_ES"
  /note="Oligo dT primed, full-length enriched cDNA library
  from undifferentiated hES cell lines H1 (p32), H7 (p29),
  and H9 (p26) maintained in feeder-free conditions"

Query Match
Best Local Similarity 100.0%; Pred. No. 2.4e-256; Length 704;
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

25 GCGCGTGTACTGCGCGCTGCTTCTTAGTCGCGGTTCCGCCGCTGGTGTGTCACCG 84
3 GCGCGTGTACTGCGCGCTGCTTCTTAGTCGCGGTTCCGCCGCTGGTGTGTCACCG 62
85 GCGCGTGTACTGCGCGCTGCTTCTTAGTCGCGGTTCCGCCGCTGGTGTGTCACCG 144
63 GCGCGTGTACTGCGCGCTGCTTCTTAGTCGCGGTTCCGCCGCTGGTGTGTCACCG 122
145 GCGCGTGTACTGCGCGCTGCTTCTTAGTCGCGGTTCCGCCGCTGGTGTGTCACCG 204
123 GCGCGTGTACTGCGCGCTGCTTCTTAGTCGCGGTTCCGCCGCTGGTGTGTCACCG 182
205 AAATGCAACATGGCGAGCAATGGAACAGACAGCTGGTGTGAGATATTTGAAACT 264

```

[5'-TGTTACCAATCTGAAGTGGAGCGCGCTCATTTTCTTTTCTTTT-3',
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Donaldso."

ORIGIN	
Query Match	15.9%; Score 495; DB 1; Length 500;
Best Local Similarity	100.0%; Pred. No. 81e-256;
Matches 495; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	2625 TGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGATGAGATGTCATTTTGACCA 2684
Db	495 TGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGATGAGATGTCATTTTGACCA 436
QY	2685 TACCAAGTATTTACTGAGTTTCTTGTAGTGGGCTGGAAAGCCATATGATTTACAGATCTA 2744
Db	435 TACCAAGTATTTACTGAGTTTCTTGTAGTGGGCTGGAAAGCCATATGATTTACAGATCTA 376
QY	2745 TCCTCAGGAGACACAGCATAAGATTCCTGAATCGGGAGAACATTTATGAATCGATCT 2804
Db	375 TCCTCAGGAGACACAGCATAAGATTCCTGAATCGGGAGAACATTTATGAATCGATCT 316
QY	2805 TTTCACACTACCTTCAAGAAACCTTGGATCACGTTTGTCTCTAAAGTGCATTAAT 2864
Db	315 TTTCACACTACCTTCAAGAAACCTTGGATCACGTTTGTCTCTAAAGTGCATTAAT 256
QY	2865 TTGACCTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAATGAGAGGTTTAAT 2924
Db	255 TTGACCTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAATGAGAGGTTTAAT 196
QY	2925 CAACAGAAACACAGAAATGATCATCATTTTGTATACCTGCGATGTAACATCTACTCCT 2984
Db	195 CAACAGAAACACAGAAATGATCATCATTTTGTATACCTGCGATGTAACATCTACTCCT 136
QY	2985 GAAATAAATGTGTGTCATGTCAGGGGTCTACGGTTTGTGGTAACTATACCTTAA 3044
Db	135 GAAATAAATGTGTGTCATGTCAGGGGTCTACGGTTTGTGGTAACTATACCTTAA 76
QY	3045 CCCACATGCTCAAAATCAATGATACATATTCCTGAGAGCCAGCAATACCATAGAA 3104
Db	75 CCCACATGCTCAAAATCAATGATACATATTCCTGAGAGCCAGCAATACCATAGAA 16
QY	3105 TTACTAAAAA 3119
Db	15 TTACTAAAAA 1

RESULT 110
BE789604
LOCUS 601481552F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3884197 5',
DEFINITION mRNA sequence.
ACCESSION BE789604
VERSION BE789604.1 GI:10210802
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 617)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Contact: Ince Genomics, Inc.
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM9657 row: 9 column: 14
High quality sequence stop: 614.
Location/Qualifiers
1. 617
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3884197"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_68"
/note="Organ: lung; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.8 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match	15.8%; Score 493; DB 7; Length 617;
Best Local Similarity	99.8%; Pred. No. 9.9e-255;
Matches 613; Conservative	0; Mismatches 0; Indels 1; Gaps 1;
QY	2213 GAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTGGTTGTAGTCATAG 2272
Db	1 GAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTGGTTGTAGTCATAG 60
QY	2273 ACAACAGGGATCTCTGTACCGAGGGCTTAAATTTGAAGCGCTTTTAAATATAAAATGG 2332
Db	61 ACAACAGGGATCTCTGTACCGAGGGCTTAAATTTGAAGCGCTTTTAAATATAAAATGG 120
QY	2333 GTCAA -ATAGAAATTTGACGATCAGGTGGAAGACTCCAATATCTAGCTTCTCGATATGAT 2391
Db	121 GTCAAATAGAAATTTGACGATCAGGTGGAAGACTCCAATATCTAGCTTCTCGATATGAT 180
QY	2392 TTCAATTGACTTGTAGTATCGTGTGGCATCCAGGCTGGTCTATGGAGTACCTCTCCCTG 2451
Db	181 TTCAATTGACTTGTAGTATCGTGTGGCATCCAGGCTGGTCTATGGAGTACCTCTCCCTG 240
QY	2452 ATGCATTAATGACAGAGTCAAGATATCTTCAAGGTTGTCTATGCTGGGGCCCACTCACT 2511
Db	241 ATGCATTAATGACAGAGTCAAGATATCTTCAAGGTTGTCTATGCTGGGGCCCACTCACT 300
QY	2512 CTGTGGATCTTCTATGATACAGGATACAGGAACCTTATGCGGTACACCTGACCAAGAT 2571
Db	301 CTGTGGATCTTCTATGATACAGGATACAGGAACCTTATGCGGTACACCTGACCAAGAT 360
QY	2572 GAACAGGGCTATTACTTAGGATCTGGGCATCGAAGCAGAAAGTTCCCTCTGAACCA 2631
Db	361 GAACAGGGCTATTACTTAGGATCTGGGCATCGAAGCAGAAAGTTCCCTCTGAACCA 420
QY	2632 AATCGTTTACTGCTCTTACATCGTTTCTGATGAGATGTCATTTTGGACATACCAGT 2691
Db	421 AATCGTTTACTGCTCTTACATCGTTTCTGATGAGATGTCATTTTGGACATACCAGT 480
QY	2692 ATATTACTGAGTTTTTGTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCCTCAG 2751
Db	481 ATATTACTGAGTTTTTGTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCCTCAG 540
QY	2752 GAGAGACACAGCATAGAGTTCCTGAAATCGGGAGAACATTTATGAATCTGTCATCTTTTCAC 2811
Db	541 GAGAGACACAGCATAGAGTTCCTGAAATCGGGAGAACATTTATGAATCTGTCATCTTTTCAC 600
QY	2812 TACCTTCAAGAAA 2825
Db	601 TACCTTCAAGAAA 614

RESULT 111
DA720650
LOCUS DA720650 NT2R13 Homo sapiens cDNA clone NT2R13000840 5', mRNA
DEFINITION linear EST 13-NOV-2005
sequence.

ACCESSION DA720650
 VERSION DA720650.1 GI:82366371
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 REFERENCE 1 (bases 1 to 492)
 AUTHORS Kimura.K., Wakamatsu.A., Suzuki.Y., Ota.T., Nishikawa.T.,
 Yamashita.R., Yamamoto.J., Sekine.M., Tsuritani.K., Wakaguri.H.,
 Ishii.S., Sugiyama.T., Saito.K., Isono.Y., Irie.R., Kushida.N.,
 Yoneyama.T., Otsuka.R., Kanda.K., Yokoi.T., Kondo.H., Wagatsuma.M.,
 Murakawa.K., Ishida.S., Iehibashi.T., Takahashi-Fujii.A.,
 Tanase.T., Nagai.K., Kikuchi.H., Nakai.K., Isogai.T. and Sugano.S.
 TITLE Diversification of Transcriptional Modulation: Large-scale
 Identification and Characterization of Putative Alternative
 Promoters of Human Genes
 JOURNAL Genome Res. 16 (1), 55-65 (2006)
 PubMed 16344560
 COMMENT Contact: Takao Isogai
 FLJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Kamatari, Kisearazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: flj-cdna@nifty.com
 NEDO human cDNA project (New Energy and Industrial Technology
 Developmental Organization, Japan); cDNA library construction:
 Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
 Research Association for Biotechnology (RAB) and Biotechnology
 Center, National Institute of Technology and Evaluation; 3'-end one
 pass sequencing: RAB.
 FEATURES Location/Qualifiers
 1..492
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NT2R13000840"
 /cell_type="teratocarcinoma"
 /cell_line="NT2"
 /clone_lib="NT2R13"
 /note="Vector: pME18SFL3; majorly NT2 neuron; mRNA from
 NT2 neuronal precursor cells treated 2-weeks mitotic
 inhibitor after 5-weeks retinoic acid (RA) induction"
 ORIGIN
 Query Match 15.8%; Score 492; DB 9; Length 492;
 Best Local Similarity 100.0%; Pred. No. 3.4e-254;
 Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAGTGCTAAAGCCTCCGAGGCCGAGCGCGTCTACTGCGCGCGCTGCTTCTTAGTGGCG 60
 DB 1 AAGTGCTAAAGCCTCCGAGGCCGAGCGCGTCTACTGCGCGCGCTGCTTCTTAGTGGCG 60
 QY 61 CGTTGCGCGCTGGGTTGTCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 DB 61 CGTTGCGCGCTGGGTTGTCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 QY 121 GAGTGAGCGCGCGCGAGCATGAAGCGCGCGCGCGCGCGCGCTCCATAGCGCAGTCCGAGCG 180
 DB 121 GAGTGAGCGCGCGCGAGCATGAAGCGCGCGCGCGCGCGCGCTCCATAGCGCAGTCCGAGCG 180
 QY 181 TCCGCGCGCGCGCGCGGGAAGAAATGCAACATGCGAGCAGCAATGGAACAGAACAG 240
 DB 181 TCCGCGCGCGCGCGCGGGAAGAAATGCAACATGCGAGCAGCAATGGAACAGAACAG 240
 QY 241 CTGGGTGTTGAGATATTTGAACTCGGACTGTGAGGAGATATTTGAATCAGAGGATCG 300
 DB 241 CTGGGTGTTGAGATATTTGAACTCGGACTGTGAGGAGATATTTGAATCAGAGGATCG 300
 QY 301 CCTAAATTTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT 360

Db 301 CCTAAATTTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT 360
 QY 361 GCCGATACACGAAAATATCATGGCTACATGATGGCTAAGGACGACCATGATTTTCATGTTT 420
 Db 361 GCCGATACACGAAAATATCATGGCTACATGATGGCTAAGGACGACCATGATTTTCATGTTT 420
 QY 421 GTGAAGAGGATGATCCAGTGGACCTATTTCAGACAAATCTATTACCTTGGCCATGCT 480
 Db 421 GTGAAGAGGATGATCCAGTGGACCTATTTCAGACAAATCTATTACCTTGGCCATGCT 480
 QY 481 GGTGAGAGACAGA 492
 Db 481 GGTGAGAGACAGA 492
 RESULT 112
 BGI65205
 LOCUS 60234407AF1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4453987 5',
 DEFINITION mRNA sequence.
 ACCESSION BGI65205
 VERSION BGI65205.1 GI:12671908
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 REFERENCE 1 (bases 1 to 639)
 AUTHORS NIH-MGC Htp://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: sgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 High quality sequence stop: 639.
 Location/Qualifiers
 1..639
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4453987"
 /tissue_type="hypernephroma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 89"
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC library."
 ORIGIN
 Query Match 15.5%; Score 485; DB 2; Length 639;
 Best Local Similarity 99.8%; Pred. No. 2.2e-250;
 Matches 605; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1878 CTCACATCTTGTGATCAGTCAGTCAGTCAGTCCTTCTTATAGCTAGTATAGTAACCA 1937
 Db 1 CTCACATCTTGTGATCAGTCAGTCAGTCAGTCCTTCTTATAGCTAGTATAGTAACCA 60
 QY 1938 GAAGAATCCACATCTGTGTGTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTG 1997
 Db 61 GAAGAATCCACATCTGTGTGTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTG 120
 QY 1998 CAAAACAAAGAAATTTGGGCGCACCAATTTTGATTCAGCAGGTCTCTCTCGACTATAC 2057

Db	121	CAAAACAAAGGAATTTGGGCCACCA-ITTTGGATTACAGAGGTCTCTTCTCGACTATAC	179
Qy	2058	TCCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTGTATGGGATGCTCTA	2117
Db	180	TCCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTGTATGGGATGCTCTA	239
Qy	2118	CAAGCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGCTGTTTCATATATGGTGG	2177
Db	240	CAAGCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGCTGTTTCATATATGGTGG	299
Qy	2178	TCCTCAGGTGCAGTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATAC	2237
Db	300	TCTCTCAGGTGCAGTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATAC	359
Qy	2238	CCTACGCTCTCTAGGTTATGTGGTTGTAGTGATAGACACAGGGGATCCTGTCAACGAGG	2297
Db	360	CCTACGCTCTCTAGGTTATGTGGTTGTAGTGATAGACACAGGGGATCCTGTCAACGAGG	419
Qy	2298	GCTTAAATTTTCAAGGCGCCTTTAAATATATAAATGGGTCAATAGAAATTTGACAGTCAGGT	2357
Db	420	GCTTAAATTTGAGGCGCCTTTAAATATATAAATGGGTCAATAGAAATTTGACAGTCAGGT	479
Qy	2358	GGAGGACTCAATATCTAGCTTCTCGATATGATTTCAATTGACTTAGATCGTGTGGGCAT	2417
Db	480	GGAGGACTCAATATCTAGCTTCTCGATATGATTTCAATTGACTTAGATCGTGTGGGCAT	539
Qy	2418	CCAAGGTGCTCTATGGAGGATACCTCTCCCTGATGGCATTTAATGACAGAGGTCAAGTAT	2477
Db	540	CCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCATTTAATGACAGAGGTCAAGTAT	599
Qy	2478	CTTCAG 2483	
Db	600	CTTCAG 605	

[illegible]

REFERENCE	
AUTHORS	Hominidae; Homo. 1 (bases 1 to 553) Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakisugi, H., Ishii, S., Sugiyama, T., Saito, K., Iseno, Y., Irie, R., Kunita, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fuji, A., Waga, S., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isozaki, T. and Sugano, S.
TITLE	Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
JOURNAL	Genome Res. 16 (1), 55-65 (2006)

Contact: Takao Isogai
FLJ Project (HRI team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com

NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction,
Helix Research Institute (HRI); 5'-end one pass sequencing; HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing; RAB.

FEATURES	source	Location/Qualifiers
		1..553
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/clone="FEBRA2012626"
		/tissue_type="brain"
		/dev_stages="fetal"
		/clone_lib="FEBRA2"
		/note="Vector: pME18SFL3"
ORIGIN		
Query Match	15.4%;	Score 481; DB 9; Length 553;
Best Local Similarity	99.8%;	Pred. No. 3.2e-248;
Matches 531;	Conservative 0;	Mismatches 1; Indels 0; Gaps 0;
QY	54	AGTGC CGGTTGCGCGCGCTGGTTGTACCGCGCCCGCCCGGAGGCAATGCAATGCGAGCGCAATGCAAC 113
DB	1	AGTGC CGGTTGCGCGCGCTGGTTGTACCGCGCGCCCGCCCGGAGGCAATGCAATGCGAGCGCAATGCAAC 60
QY	114	AGGACCGGATGGAGGCGCGCGAGCATGAAGCGCGCAGCGCCGCTCCATAGCGCACGTC 173
DB	61	AGGACCGGATGGAGGCGCGCGAGCATGAAGCGCGCAGCGCCGCTCCATAGCGCACGTC 120
QY	174	GGGACGGTCCGGCGCGCGCGCGGGGGAAGAAAATGCAATGCGAGCGCAATGCAAC 233
DB	121	GGGACGGTCCGGCGCGCGCGCGGGGGAAGAAAATGCAATGCGAGCGCAATGCAAC 180
QY	234	AGAACAGCTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCACA 293
DB	181	AGAACAGCTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCACA 240
QY	294	GGATCGGCCTAAATTTGGAGCCTTTTATGTGAGCGGTATTCCTGGAGTCACTTTAAATA 353
DB	241	GGATCGGCCTAAATTTGGAGCCTTTTATGTGAGCGGTATTCCTGGAGTCACTTTAAATA 300
QY	354	GCTGCTTGGCGATACACAGAAAATATCATGCTACATGATGGCTAGGACCAATGATTT 413
DB	301	GCTGCTTGGCGATACACAGAAAATATCATGCTACATGATGGCTAGGACCAATGATTT 360
QY	414	CATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTCAGACAAATCTATTACCTTGC 473
DB	361	CATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTCAGACAAATCTATTACCTTGC 420
QY	474	CATGCTCGTGAGACAGAGAAAATACACTGTTTATTCGAAATTCCTCAAACTATCAA 533
DB	421	CATGCTCGTGAGACAGAGAAAATACACTGTTTATTCGAAATTCCTCAAACTATCAA 480
QY	534	TAGACAGCAGCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTCAG 585
DB	481	TAGACAGCAGCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTCAG 532
RESULT 114		
LOCUS	BI710784	
DEFINITION	489008.y1 Human insulinoma Homo sapiens mRNA	536 bp linear EST 11-MAR-2002
ACCESSION	5, similar to TR:075273 O75273 R26984_1	's, mRNA sequence.
VERSION	BI710784.1	
KEYWORDS	EST:	GI:15686479
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	
	Hominidae; Homo.	
	1 (bases 1 to 536)	
	Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,	
	Lewinshka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,	
	Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,	
	Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,	
	Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R.,	
	Williams,T., Jackson,Y. and Bowers,Y.	

Thu Jun 22 09:04:26 2006

```

TITLE      Endocrine Pancreas Consortium
JOURNAL    Unpublished (2000)
COMMENT    Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
          Endocrine Pancreas Consortium
          Harvard University, Howard Hughes Medical Institute
          Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
          MA 02138
          Tel: 617-495-1812
          Fax: 617-495-8557
          Email: dmelton@biohp.harvard.edu
          Library was constructed by Dr. J. Ferrer In vivo mass-excised to
          pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
          University Genome Sequencing Center For information on obtaining a
          clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
          Seq primer: -40RP from Gibco
          High quality sequence stop: 428.

FEATURES   Location/Qualifiers
            1..536
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:5023286"
             /tissue_type="insulinoma"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="Human insulinoma"
             /note="Organ: pancreas; Vector: pBluescript SK-; Site:1:
             XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system
             (stratagene) by Dr. J. Ferrer, in vivo mass-excised to
             pBluescript SK- by Dr. H. Inoue following the Washington
             University protocol
             (http://genome.wustl.edu/est/lambda_protocol.shtml).
             Please contact Hiroshi Inoue, MD/PhD for further
             information on this library [Metabolism Division, Permutt
             Laboratory, Washington University School of Medicine, Box
             8127, 660 S Euclid Ave, St. Louis, MO 63110]. Note: this
             is a Washington University Pancreas EST project library."

ORIGIN
Query Match      15.4%; Score 480; DB 2; Length 536;
Best Local Similarity 99.8%; Pred. No. 1.1e-247;
Matches 530; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2532 AGGATACAGGACGTTATATGGTCCACCTGACAGCAATGAACAGGCTATTACTTAGG 2591
Db 6 AGGATACAGGACGTTATATGGTCCACCTGACAGCAATGAACAGGCTATTACTTAGG 65
Qy 2592 ATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAACCAATGTTTACTGCTTTACA 2651
Db 66 ATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAACCAATGTTTACTGCTTTACA 125
Qy 2652 TGGTTTCTGGATGAGATGTCCTATTTTGACATACCAGTATATTACTGAGTTTGTAGT 2711
Db 126 TGGTTTCTGGATGAGATGTCCTATTTTGACATACCAGTATATTACTGAGTTTGTAGT 185
Qy 2712 GAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCAATGAAGT 2771
Db 186 GAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCAATGAAGT 245
Qy 2772 TCCTGAATCGGGAGAACATTATGAACTGTCATCTTTTGCACTACCTTCAGAAACCTTGG 2831
Db 246 TCCTGAATCGGGAGAACATTATGAACTGTCATCTTTTGCACTACCTTCAGAAACCTTGG 305
Qy 2832 ATCAGTATGCTGCTCTTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATA 2891
Db 306 ATCAGTATGCTGCTCTTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATA 365
Qy 2892 CACTGGCTATTTAAACAAATGAGGAGGTTTAAATCAACAGAAACACAGAAATTCATCA 2951
Db 366 CACTGGCTATTTAAACAAATGAGGAGGTTTAAATCAACAGAAACACAGAAATTCATCA 425
Qy 2952 CATTTTGATACCTGCCATGTAACATCTACTCTCGAAATAAATGTTGGTGCATGCGAGGG 3011
Db 426 CATTTTGATACCTGCCATGTAACATCTACTCTCGAAATAAATGTTGGTGCATGCGAGGG 485

```

```

Qy 3012 TCTACGTTTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAATC 3062
Db 486 TCTACGTTTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAATC 536

RESULT 115
LOCUS   BI084885 855 bp mRNA linear EST 20-JUN-2001
DEFINITION 602869453T1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5013996 3',
          mRNA sequence.
ACCESSION BI084885
VERSION   BI084885.1 GI:14503215
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
           Hominidae; Homo.
REFERENCE 1 (bases 1 to 855)
AUTHORS  NIH-MGC http://mgi.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: CLONTECH Laboratories, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLC1820 row: j column: 13
          High quality sequence start: 3
          High quality sequence stop: 844.
FEATURES   Location/Qualifiers
            1..855
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:5013996"
             /tissue_type="epidermoid carcinoma, cell line"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="NIH_MGC_102"
             /note="Organ: salivary gland; Vector: pOTB7; Site:1: XhoI;
             Site 2: EcoRI; cDNA made by oligo-dT priming
             Directionally cloned into EcoRI/XhoI sites using the
             following 5' adaptor: GGCACGAG(G). Library constructed
             by Ling Hong in the laboratory of Gerald M. Rubin
             (University of California, Berkeley) using ZAP-cDNA
             synthesis kit (Stratagene) and Superscript II RT (Life
             Technologies). Note: this is a NIH_MGC Library."

ORIGIN
Query Match      15.4%; Score 479; DB 2; Length 855;
Best Local Similarity 100.0%; Pred. No. 4e-247;
Matches 479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2462 TGCAGAGGTTCAGATATCTTCAGGGTTCCTATTCCTGGGGCCCCCAGTCCTCTGATCT 2521
Db 493 TGCAGAGGTTCAGATATCTTCAGGGTTCCTATTCCTGGGGCCCCCAGTCCTCTGATCT 434
Qy 2522 TCTATGATACAGGATACACGGAACTTATATGGTTCACCTTCAGCAAGATGAACAGGCT 2581
Db 433 TCTATGATACAGGATACACGGAACTTATATGGTTCACCTTCAGCAAGATGAACAGGCT 374
Qy 2582 ATTACTTACGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAATGTTTAC 2641
Db 373 ATTACTTACGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAATGTTTAC 314
Qy 2642 TGCTCTTACATGGTTTCTCGGATGAGATGTCCTATTTTGCACATACAGATATATTACTGA 2701
Db 313 TGCTCTTACATGGTTTCTCGGATGAGATGTCCTATTTTGCACATACAGATATATTACTGA 254

```


tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match	15.3%;	Score	476;	DB	1;	Length	577;
Best Local Similarity	100.0%;	Pred. NO.	1.6e-245;				
Matches	476;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						

QY	2634	TGCTTACTGCTCTTACATGTTTCTCGATGAGATGTCCATTTTGGACATACACAGTAT	2693
DB	476	TGCTTACTGCTCTTACATGTTTCTCGATGAGATGTCCATTTTGGACATACACAGTAT	417
QY	2694	ATTACTGAGTTTTTTTGTAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCCACGA	2753
DB	416	ATTACTGAGTTTTTTTGTAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCCACGA	357
QY	2754	GAGACACAGCATACAGTTCTGATCGGAGACATATGAACTGCATCTTTTGCACTA	2813
DB	356	GAGACACAGCATACAGTTCTGATCGGAGACATATGAACTGCATCTTTTGCACTA	297
QY	2814	CCTTCAAGAAAACCTTGGATCACGTTATGCTCTTAAAGTGAATATTTGACCTGT	2873
DB	296	CCTTCAAGAAAACCTTGGATCACGTTATGCTCTTAAAGTGAATATTTGACCTGT	237
QY	2874	GTAGAACTCTCTGGTATACATGCTGTTTAAACCAATGAGGAGTTTATCAACAGAA	2933
DB	236	GTAGAACTCTCTGGTATACATGCTGTTTAAACCAATGAGGAGTTTATCAACAGAA	177
QY	2934	ACAGAAATGATCATCATTTTGCATCTGACCTGCAATGTAACATCTACTCTGAAAATAA	2993
DB	176	ACAGAAATGATCATCATTTTGCATCTGACCTGCAATGTAACATCTACTCTGAAAATAA	117
QY	2994	TGTGGTGCCATGACAGGGGCTACGGTTTGTGTAGTAAATCTAAATACCTTAAACCCACATG	3053
DB	116	TGTGGTGCCATGACAGGGGCTACGGTTTGTGTAGTAAATCTAAATACCTTAAACCCACATG	57
QY	3054	CTCAAAATCAATGATACATATTCCTGAGAGAGCCAGCAATACCAATGAATTAAT	3109
DB	56	CTCAAAATCAATGATACATATTCCTGAGAGAGCCAGCAATACCAATGAATTAAT	1

RESULT 118

DA798581
LOCUS
DEFINITION
DA798581 OCBF3 Homo sapiens cDNA clone OCBF3002013 5', mRNA
sequence.

ACCESSION
DA798581.1 GI:81236986

VERSION
EST.

KEYWORDS
Homo sapiens (human)

SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushiida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Iehibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560

JOURNAL

PUBLISHED
Contact: Takao Isogai

COMMENT
FLJ Project (HRI Team)

Helix Research Institute

2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: flj-cdna@nifty.com

NEDO human cDNA project (New Energy and Industrial Technology

Developmental Organization, Japan); cDNA library construction:

Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,

Research Association for Biotechnology (RAB) and Biotechnology

Center, National Institute of Technology and Evaluation; 3'-end one

pass sequencing: RAB.

Location/Qualifiers

1. 575

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="OCBBF3002013"

/tissue_type="brain"

/dev_stage="fetal"

/clone_lib="OCBBF3"

/note="Vector: pME18SFL3"

ORIGIN

Query Match	15.1%;	Score	472;	DB	9;	Length	575;
Best Local Similarity	99.7%;	Pred. No.	2.4e-243;				
Matches	572;	Conservative	0;	Mismatches	2;	Indels	0;
Gaps	0;						

QY	8	AAAGCTCCGAGCCCAAGCCGCTGCTACTGCGCGCGCTCTTCTTAGTGGCGGTTCCG	67
DB	2	AAAGCTCCGAGCCCAAGCCGCTGCTACTGCGCGCGCTCTTCTTAGTGGCGGTTCCG	61
QY	68	CGCTGGTGGTTGTACCGCGCGCGCGCGAGGAGCCACTGCAACAGGACCGAGTGA	127
DB	62	CGCTGGTGGTTGTACCGCGCGCGCGCGAGGAGCCACTGCAACAGGACCGAGTGA	121
QY	128	GGCGGCGCAGCATGAAGCGCGCGCGCGCTTAGCGCAGCTGCGGAGCGTCCGCGC	187
DB	122	GGCGGCGCAGCATGAAGCGCGCGCGCGCTTAGCGCAGCTGCGGAGCGTCCGCGC	181
QY	188	GGGCGCGGCGGAGGAAATGCAATGGCAGCAGCAATGAAAACAGACAGCTGGTG	247
DB	182	GGGCGCGGCGGAGGAAATGCAATGGCAGCAGCAATGAAAACAGACAGCTGGTG	241
QY	248	TTGAGATATTTGAAACTGCGGACTGTGAGGAGATATTTGAATCAGAGATCGGCTAAAT	307
DB	242	TTGAGATATTTGAAACTGCGGACTGTGAGGAGATATTTGAATCAGAGATCGGCTAAAT	301
QY	308	TGGAGCCTTTTATGTTGAGCGGTATTTCTTGGAGTCACTTAAAAAGCTGTTGCCGATA	367
DB	302	TGGAGCCTTTTATGTTGAGCGGTATTTCTTGGGTGAGCTTAAAAAGCTGTTGCCGATA	361
QY	368	CCAGAAAATATCATGGCTTACATGATGGCTAAGGACCAACATGATTTTCATGTTGTGAAGA	427
DB	362	CCAGAAAATATCATGGCTTACATGATGGCTAAGGACCAACATGATTTTCATGTTGTGAAGA	421
QY	428	GGAATGATCCAGATGGACCTCAATTCACAGAGATCTATTACCTTGGCATGCTCTGGTGAGA	487
DB	422	GGAATGATCCAGATGGACCTCAATTCACAGAGATCTATTACCTTGGCATGCTCTGGTGAGA	481
QY	488	ACAGAGAAAATACACTGTTTATTTCTGAAATTTCCAAAACCTATCAATAGAGAGAGCTCT	547
DB	482	ACAGAGAAAATACACTGTTTATTTCTGAAATTTCCAAAACCTATCAATAGAGAGAGCTCT	541
QY	548	TAATGCTCTCTTTGGAGCCTCTTTTGGATCTTTT	581
DB	542	TAATGCTCTCTTTGGAGCCTCTTTTGGATCTTTT	575

RESULT 119

DA189698

LOCUS

DEFINITION

DA189698 BRAMY3 Homo sapiens cDNA clone BRAMY3011692 5', mRNA

sequence.

ACCESSION

DA189698

Accession

```

VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
Kimura,K., Wakamatsu,A., Suzuki,Y., Oka,T., Nishikawa,T.,
Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
Iehi,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kishida,N.,
Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genomics Res. 16 (1), 55-65 (2006)

JOURNAL
PUBMED
16344560
COMMENT
Contact: Takao Isogai
PLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdn@infity.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library constructing:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
Location/Qualifiers
1. .521
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRAMY3011692"
/tissue_type="amygdala"
/clone_lib="BRAMY3"
/notes="Vector: pME18SFL3"

ORIGIN
Query Match 15.0%; Score 468; DB 9; Length 521;
Best Local Similarity 99.8%; Pred No. 3.6e-241;
Matches 518; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAGTGCTAAGCCTCGAGGCCAGCGCGCTGCTACTGCGCGCGCTGCTTTAGTGGCG 60
Db 3 AAGTGCTAAGCCTCGAGGCCAGCGCGCTGCTACTGCGCGCGCTGCTTTAGTGGCG 62
QY 61 CGTTGCGCGCTGGTGTGTACCGCGCGCGCGCGAGGAGCACTGCAACAGGACCG 120
Db 63 CGTTGCGCGCTGGTGTGTACCGCGCGCGCGCGAGGAGCACTGCAACAGGACCG 122
QY 121 GAGTGGAGCGCGGAGCATGAGCGCGCGCGCGCGCTCATAGCGCAGCTGCGGACCG 180
Db 123 GAGTGGAGCGCGGAGCATGAGCGCGCGCGCGCGCTCATAGCGCAGCTGCGGACCG 182
QY 181 TCGGGCGCGCGCGGGGGAGAGGAAATGCAACATGCGCAGCAGCAATGAAACAGAACAG 240
Db 183 TCGGGCGCGCGCGGGGGAGAGGAAATGCAACATGCGCAGCAGCAATGAAACAGAACAG 242
QY 241 CTGGGTGTGAGATATTTGAAACTCGGACTGTGAGGAGAAATTTGAAATCAGAGTGG 300
Db 243 CTGGGTGTGAGATATTTGAAACTCGGACTGTGAGGAGAAATTTGAAATCAGAGTGG 302
QY 301 CCTAAATTGGAGCCCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAGCTGTT 360
Db 303 CCTAAATTGGAGCCCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAGCTGTT 362
QY 361 GCCGATACAGAAATATCATGGCTACATGATGGCTAAGGCACCATGATTCATGTT 420

```

```

Db 363 GCCGATACAGAAATATCATGGCTACATGATGCTAAGGCACCATGATTCATGTT 422
QY 421 GTGAAGAGGATGATCCAGATGGACCTCATTCACAGAGATCTATACCTTGCATGCT 480
Db 423 GTGAAGAGGATGATCCAGATGGACCTCATTCACAGAGATCTATACCTTGCATGCT 482
QY 481 GGTGAGAACAGAGAAATACACTGTTTATTCCTGAAATT 519
Db 483 GGTGAGAACAGAGAAATACACTGTTTATTCCTGAAATT 521

RESULT 120
LOCUS AA306656/c
DEFINITION EST177562 Jurkat T-cells VI Homo sapiens cDNA 5' end, mRNA
ACCSSION AA306656
VERSION AA306656.1 GI:1959055
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 457)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Buit,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.Jr.,
Glocke,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palanco,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., J.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weisner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Olsen,H.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
7566098
Other ESTs: THC184894
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M3 Reverse.
FEATURES
Location/Qualifiers
1. .457
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):160827"
/db_xref="taxon:9606"
/cell_type="T-lymphocyte"
/clone_lib="Jurkat T-cells VI"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Query Match 14.6%; Score 457; DB 1; Length 457;
Best Local Similarity 100.0%; Pred. No. 3.3e-235;
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
QY 2658 CCTGGATGAGAAATGTCATTTTGACATACCACTATATTAAGTCTTTAGTGAGGCG 2717
Db 457 CCTGGATGAGAAATGTCATTTTGACATACCACTATATTAAGTCTTTAGTGAGGCG 398
QY 2718 TGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAGAGTTCCTGA 2777
Db 397 TGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAGAGTTCCTGA 338
QY 2778 ATCGGAGAACATTAATGAATGTCATCTTTTGCACTACCTTTCAAGAAAACCTTTGGATCAG 2837
Db 337 ATCGGAGAACATTAATGAATGTCATCTTTTGCACTACCTTTCAAGAAAACCTTTGGATCAG 278
QY 2838 TATTGCTGCTCTAAAGTGATTAATTTTGACCTGTGTAGAACTCTCTGGGTATACACTGG 2897
Db 277 TATTGCTGCTCTAAAGTGATTAATTTTGACCTGTGTAGAACTCTCTGGGTATACACTGG 218
QY 2898 CTATTTAAACCAATGAGGAGGTTTAATCAACAGAAAACACAGAAATGATCATCATTTT 2957
Db 217 CTATTTAAACCAATGAGGAGGTTTAATCAACAGAAAACACAGAAATGATCATCATTTT 158
QY 2958 GATACCTGCCATGTAACATCTACTCTCGAAAATAAATGTGGTCCCATGACAGGGGTCTACG 3017
Db 157 GATACCTGCCATGTAACATCTACTCTCGAAAATAAATGTGGTCCCATGACAGGGGTCTACG 98
QY 3018 GTTTGTGTAGTAACTTAATACCTTAACCCCATCTCAAAATCAAAATGATACATATTC 3077
Db 97 GTTTGTGTAGTAACTTAATACCTTAACCCCATCTCAAAATCAAAATGATACATATTC 38
QY 3078 CTGAGAGACCCAGCAATACCATAAGAAATTAATAAAA 3114
Db 37 CTGAGAGACCCAGCAATACCATAAGAAATTAATAAAA 1
```

RESULT 121

```
DB156905 THYM3 Homo sapiens 576 bp mRNA linear EST 04-DEC-2005
LOCUS DB156905 THYM3 Homo sapiens cDNA clone THYM3035791 5', mRNA
DEFINITION sequence.
```

ACCESSION DB156905

VERSION DB156905.1 GI:83187534

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 576)

Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S. Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes

Genome Res. 16 (1), 55-65 (2006)

16344560

Contact: Takao Isogai

FLJ Project (HRI Team)

Helix Research Institute

2-6-7 Karusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: flj-cdna@nifty.com

NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction; Helix Research Institute (HRI); 5'-end one pass sequencing; HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing; RAB.

FEATURES

Location/Qualifiers

source

```
1..576
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="THYM3035791"
/tissue_type="thymus"
/clone_lib="THYM3"
/note="Vector: pME18SFL3"
```

ORIGIN

```
Query Match 14.6%; Score 455; DB 9; Length 576;
Best Local Similarity 99.8%; Pred. No. 4e-234;
Matches 575; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AAGTGTCAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCCCTGCTTCTTAGTGC 60
Db 1 AAGTGTCAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCCCTGCTTCTTAGTGC 60
QY 61 CGTTGCCGCCCTGGGTTGTCAACCGCGCGCGCGGAGGAGCCACTGCAACACGAGAC 120
Db 61 CGTTGCCGCCCTGGGTTGTCAACCGCGCGCGCGGAGGAGCCACTGCAACACGAGAC 120
QY 121 GAGTGAGCGCGCGCAGCATGAAGC-GGCGCAGCGCGCTCCATAGCGCAGCTCGGAG 179
Db 121 GAGTGAGCGCGCGCAGCATGAAGC-GGCGCAGCGCGCTCCATAGCGCAGCTCGGAG 180
QY 180 GTCCGCGCGCGCGCGGAGGAGAAATGCAACATGCGCAGCAGCAATGGAACAGACA 239
Db 181 GTCCGCGCGCGCGCGGAGGAGAAATGCAACATGCGCAGCAGCAATGGAACAGACA 240
QY 240 GCTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACA 299
Db 241 GCTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACA 300
QY 300 GCCTAAATGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCT 359
Db 301 GCCTAAATGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCT 360
QY 360 TGCCTATACACAGAAATATCATGCTACATGATGGCTAAGGCACCATGATTTTCAT 419
Db 361 TGCCTATACACAGAAATATCATGCTACATGATGGCTAAGGCACCATGATTTTCAT 420
QY 420 TGTGAAGAGGAATGATCCAGATGAGACCTCATTCAGACAGATCTTATACCTTGC 479
Db 421 TGTGAAGAGGAATGATCCAGATGAGACCTCATTCAGACAGATCTTATACCTTGC 480
QY 480 TGTGTGAGAACAGAGAAATACACTGCTTTTATTTCTGAAATTCCTCAAACTAT 539
Db 481 TGTGTGAGAACAGAGAAATACACTGCTTTTATTTCTGAAATTCCTCAAACTAT 540
QY 540 AGCAGCTTTTAATGCTCTCTTTGGAAGCCCTCTTTTGA 575
Db 541 AGCAGCTTTTAATGCTCTCTTTGGAAGCCCTCTTTTGA 576
```

RESULT 122

DA607747

LOCUS

DEFINITION

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 587)

Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,

DA607747 IMR322 Homo sapiens cDNA clone IMR3222005026 5', mRNA

DA607747 sequence.

DA607747 GI:80534653

EST. Homo sapiens (human)

DA607747 Homo sapiens

DA607747 sequence.

DA607747 GI:80534653

EST. Homo sapiens (human)

DA607747 Homo sapiens

DA607747 sequence.

DA607747 GI:80534653

EST. Homo sapiens (human)

DA607747 Homo sapiens

DA607747 sequence.

DA607747 GI:80534653

EST. Homo sapiens (human)

DA607747 Homo sapiens

DA607747 sequence.

DA607747 GI:80534653

EST. Homo sapiens (human)

[illegible][illegible]

Genome Res. 13 (2003) 1021-1024
 16344560
 Contact: Takao Isogai
 FIJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: fi_j-cdn@nifty.com
 NEDO human cDNA project (New Energy and Industrial Technology
 Developmental Organization, Japan); cDNA library construction;
 Helix Research Institute (HRI); 5'-end one pass sequencing; HRI,
 Research Association for Biotechnology (RAB) and Biotechnology
 Center, National Institute of Technology and Evaluation; 3'-end one
 pass sequencing; RAB.
 Features
 source
 1. .570
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="TRACH3030760"
 /tissue_type="trachea"
 /clone_id="TRACH3"
 /explosion="EXP.1"

ORIGIN

Query Match	14.4%;	Score 449;	DB 9;	Length 570;
Best Local Similarity	99.8%;	Pred. No. 7.2e-231;		
Matches 569;	Conservative	0;	Mismatches 21;	Indels 1; Gaps 1;
QY	8	AAAGCCTCCGAGGCCAAGGCCGCTCTACTCTGCGCCGCTGCTCTCTTAGTGC CGGTTCGC	67	
DB	1	AAAGCCTCCGAGGCCAAGGCCGCTCTACTGCCCGCGCTGCTCTTAGTGC CGGTTCGC	60	
QY	68	CGCCTGGGTGTGTCACCGCGCGCGCGCGAGGAGCCACTGCAACACGAGACCGAGTGGGA	127	
DB	61	CGCCTGGGTGTGTCACCGCGCGCGCGCGAGGAGCCACTGCAACACGAGACCGAGTGGGA	120	
QY	128	GGCGCGCGCAGCATGAAGC- GGCGCAGCGCCGCTCCATAGCGCAGCTCGGACCGTCCGGG	186	
DB	121	GGCGCGCGCAGCATGAAGCNGCGCAGGCGCGCTCCATAGCGCAGCTCGGACCGTCCGGG	180	
QY	187	CGGGCGCGGGGGNAGGAAAATGCAACATGGCAGCAGCAATGGAAAACAGACAGCTGGGT	246	
DB	181	CGGGCGCGGGGGNAGGAAAATGCAACATGGCAGCAGCAATGGAAAACAGACAGCTGGGT	240	
QY	247	GTTGAGATATTTGAAACTCGGCACTGTGAGGAGATATTGAATCACAGATCGGCTAAA	306	
DB	241	GTTGAGATATTTGAAACTCGGCACTGTGAGGAGATATTGAATCACAGATCGGCTAAA	300	
QY	307	TTGAGCGCTTTTATGCTTCGAGCGGTATTCCTCGAGTCAGCTTAAAAAGCTGCTTCCGAT	366	
DB	301	TTGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTTCCGAT	360	
QY	367	ACCAGAAAATATCATGGCTACATGATGGTGAAGGCAACCATGATTTCACTGTTGGAAG	426	
DB	361	ACCAGAAAATATCATGGCTACATGATGGTGAAGGCAACCATGATTTCACTGTTGGAAG	420	
QY	427	AGGAATGATCCAGATGAGACCTCAATCAGACAGAACTTATTACCTGCCATGCTCGGTGAG	486	
DB	421	AGGAATGATCCAGATGAGACCTCAATCAGACAGAACTTATTACCTGCCATGCTCGGTGAG	480	
QY	487	AACAGAGAAAATACACTGTTTATTTCTTGAAATTTCCCAAAAATCAATATAGCAGCAGTTC	546	
DB	481	AACAGAGAAAATACACTGTTTATTTCTTGAAATTTCCCAAAAATCAATATAGCAGCAGTTC	540	
QY	547	TTAATGCTCTCTTGGAGACCTCTTTTGGAT	576	
DB	541	TTAATGCTCTCTTGGAGACCTCTTTTGGAT	570	

RESULT 125	linear	EST 04-NOV-2005
DA082837	568 bp	568 bp
LOCUS	BRACE2 Homo sapiens cDNA clone BRACE2036520 5', mRNA	
DEFINITION	sequence.	
ACCESSION	DA082837	
VERSION	DA082837.1	GI:80467982
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 568)	
AUTHORS	Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamaashita, R., Yamamoto, J., Sekine, M., Teurittani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.	
TITLE	Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes	
JOURNAL	Genome Res. 16 (1), 55-65 (2006)	
PUBLISHED	16344560	
COMMENT	Contact: Takao Isogai EST Project (HFT Team)	

Helix Research Institute		26-7, Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan	
Tel: 81-438-52-3975			
Fax: 81-438-52-3986			
Email: flj-cdn@nifty.com			
NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction:			
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,			
Research Association for Biotechnology (RAB) and Biotechnology			
Center, National Institute of Technology and Evaluation; 3'-end one			
pass sequencing: RAB.			
FEATURES	Location/Qualifiers		
source	1..568		
	/organism="Homo sapiens"		
	/mol_type="nrna"		
	/db_xref="taxon:9606"		
	/clone="BRACE2036520"		
	/tissue_type="cerebellum"		
	/clone_lib="BRACE2"		
	/note="Vector: pME18SFL3"		
ORIGIN			
Query Match	14.3%;	Score 447;	DB 9; Length 568;
Best Local Similarity	99.8%;	Pred. No. 8.7e-230;	
Matches 567;	Conservative	0;	Mismatches 0; Indels 1; Gaps 1;
QY	1	AAAGTCTTAAGCCTCCGAGGCCAAGCGCGTGTCTACTGCCGCGCGCTCTTCTTTAGTGC	60
Db	1	AAAGTCTTAAGCCTCCGAGGCCAAGCGCGTGTCTACTGCCGCGCGCTCTTCTTTAGTGC	60
QY	61	CGTTCCCGCCTGGGTTGTCAACCGCCCGCCGCGAGGAAGCACTGCAACCGAGACG	120
Db	61	CGTTCCCGCCTGGGTTGTCAACCGCCCGCCGAGGAAGCACTGCAACCGAGACG	120
QY	121	GAGTGGAGCGCGCAGCATGAAGCGCGCAGCGCCGCTCCATAGC-GCAGCTCGGAGC	179
Db	121	GAGTGGAGCGCGCAGCATGAAGCGCGCAGCGCCGCTCCATAGCAGCTCGGAGC	180
QY	180	GTCGCGCGCGCGCGGGGGAAGAAATGCAACATGGCAGCAGCAATGGAAACAGACA	239
Db	181	GTCGCGCGCGCGCGGGGGAAGAAATGCAACATGGCAGCAGCAATGGAAACAGACA	240
QY	240	GCTGGGTGTGAGATATTGAAATCTCGGACTGTGAGGAGATATTGAATCACAGGATCG	299
Db	241	GCTGGGTGTGAGATATTGAAATCTCGGACTGTGAGGAGATATTGAATCACAGGATCG	300
QY	300	GCCTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTCT	359
Db	301	GCCTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTCT	360
QY	360	TGCCGATACAGAAAAATATCATGGCTACATGATGGCTAAGGCACCACATGATTTCA	419
Db	361	TGCCGATACAGAAAAATATCATGGCTACATGATGGCTAAGGCACCACATGATTTCA	420
QY	420	TGTGAAGAGGAATGATCCAGATGGACCTCAITCAGACAGATCTATTACCTTGCATGC	479
Db	421	TGTGAAGAGGAATGATCCAGATGGACCTCAITCAGACAGATCTATTACCTTGCATGC	480
QY	480	TGTTGAGAACACAGAAAAATACACTGTTTATTTCTGAAATCCCAAAATCAATCAGC	539
Db	481	TGTTGAGAACACAGAAAAATACACTGTTTATTTCTGAAATCCCAAAATCAATCAGC	540
QY	540	AGCAGTCTTAATGCTCTCTTGGAGCCT	567
Db	541	AGCAGTCTTAATGCTCTCTTGGAGCCT	568

DBL61177.1 GI:83155018
 EST.
 KEYWORDS
 SOURCES
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1. (bases 1 to 570)
 Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
 Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
 Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida N.,
 Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
 Morikawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
 Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
 TITLE
 Diversification of Transcriptional Modulation: Large-scale
 Identification and Characterization of Putative Alternative
 Promoters of Human Genes
 JOURNAL
 PUBLISHED
 Genome Res. 16 (1), 55-65 (2006)
 16344560
 CONTACT: Takao Isogai
 FLJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: flj-cdna@infy.com
 NEDO human cDNA project (New Energy and Industrial Technology
 Developmental Organization, Japan); cDNA library construction.
 Helix Research Institute (HRI); 5'-end one pass sequencing; HRI,
 Research Association for Biotechnology (RAB) and Biotechnology
 Center, National Institute of Technology and Evaluation, 3'-end one
 pass sequencing; RAB.
 FEATURES
 source
 Location/Qualifiers
 1..570
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="THYMU3041608"
 /issue_type="thymus"
 /clone_lib="THYMU3"
 /notes="Vector: pME18SFL3"

RESULT 126	DB161177	DB161177	570 bp	mRNA	linear	EST 03-DEC-2005
LOCUS	DB161177	DB161177	THYM3	Homo sapiens	cDNA clone	THYMU3041608 5', mRNA
DEFINITION	DB161177	DB161177	sequence.			
ACCESSION	DB161177	DB161177				


```

Db      363  TGCCGATACAGAAATATCATGGCTACATGATGGCTAAGGCACCAATGATTTTCATGTT 422
Qy      420  TGTGAAGAGAAATGATCCAGATGGACCTCATTCAGACAGATCTATTACCTTGGCAGTTC 479
Db      423  TGTGAAGAGAAATGATCCAGATGGACCTCATTCAGACAGATCTATTACCTTGGCAGTTC 482
Qy      480  TGGTGAGAACAGAGAAATACACTGTTTTATTCTGAAATTCCTCAAACTATCAATAGAGC 539
Db      483  TGGTGAGAACAGAGAAATACACTGTTTTATTCTGAAATTCCTCAAACTATCAATAGAGC 542
Qy      540  AGCAGTCTTAATGCTCTCTTGGAGGCT 567
Db      543  AGCAGTCTTAATGCTCTCTTGGAGGCT 570

RESULT 127
AA465309
LOCUS   aa24b08.t1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:814167 5',
DEFINITION
mRNA sequence.
ACCESSION
VERSION AA465309.1 GI:2191476
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 469)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham.

FEATURES
source
1. 469
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:6032116"
/db_xref="taxon:9606"
/clone="IMAGE:814167"
/tissue_type="germinal center 'B cell'"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_GCB1"
/note="Vector: p7T3D-Paci; Site 1: Not I; Site 2: Eco RI;
1st strand cDNA was prepared from human tonsillar cells
enriched for germinal center B cells by flow sorting
(CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr.
David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA
synthesis was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAATGAGGAGCGGCTCATTTTCTTTTCTTTT-3',
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 14.3%; Score 445; DB 1; Length 469;
Best Local Similarity 100.0%; Pred. No. 1e-228;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1560  GTTATTATTTTGCTCTGAATGCAGAACAGGTTTCGGTCATTTATACAAAATTACATCTAT 1619
Db      1 GTTATTATTTTGCTCTGAATGCAGAACAGGTTTCGGTCATTTATACAAAATTACATCTAT 60
Qy      1620  TTTAAAGGAAACAAATATAACAGATCCAGTGGTGGGCTGCTCTCAAGTGAATTTCAA 1679
Db      61 TTTAAAGGAAACAAATATAACAGATCCAGTGGTGGGCTGCTCTCAAGTGAATTTCAA 120

1680  GTGTCTCTCAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTTGGCCGGCA 1739
121  GTGTCTCTCAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTTGGCCGGCA 180
1740  TGGATCTTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1799
181  TGGATCTTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 240
1800  CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTAAATCTCTGGAGAGGTGACAAG 1859
241  CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTAAATCTCTGGAGAGGTGACAAG 300
1860  GCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTCAGTGTGACTTCTTTAT 1919
301  GCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTCAGTGTGACTTCTTTAT 360
1920  AAGTAAGTATAGTATACCAAGAGAAATCCACACTGTGTCTTTCACAAAGCTATCAAGTCC 1979
361  AAGTAAGTATAGTATACCAAGAGAAATCCACACTGTGTCTTTCACAAAGCTATCAAGTCC 420
1980  TGAAGATGACCCAACTTGCAGAAAACA 2004
421  TGAAGATGACCCAACTTGCAGAAAACA 445

AA831859 437 bp mRNA linear EST 18-MAR-1998
OC93h12.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1357319 3',
mRNA sequence.
AA831859
VERSION AA831859.1 GI:2904958
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 437)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1067 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 399.
Location/Qualifiers
1. 437
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1357319"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_GCB1"
/note="Vector: p7T3D-Paci; Site 1: Not I; Site 2: Eco RI;
1st strand cDNA was prepared from human tonsillar cells
enriched for germinal center B cells by flow sorting
(CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr.
David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA
synthesis was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAATGAGGAGCGGCTCATTTTCTTTTCTTTT-3',
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

```


[5'-TGTACCAATCTGAAGTGGGAGCGCGCTCATTTTTTTTTTTTTTTT-3']
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pTV73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 14.0%; Score 437; DB 1; Length 437;
Best Local Similarity 100.0%; Pred. No. 2.3e-224;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2683 CATACAGATATATTACTAGTGTCTTTAGTGGGCTGGAAGCCATATGTTTACAGATC 2742
DB CATACAGATATATTACTAGTGTCTTTAGTGGGCTGGAAGCCATATGTTTACAGATC 378
QY 2743 TATCTCTCAGGAGACACAGATAGAGTTCCTGAATCGGAGAACATTATGACATGCAAT 2802
DB TATCTCTCAGGAGACACAGATAGAGTTCCTGAATCGGAGAACATTATGACATGCAAT 318
QY 2803 CTTTTTGCACTACCTTCAGAAACCTTGGATCAGATTTGCTCTCTAAAGTGATATAA 2862
DB CTTTTTGCACTACCTTCAGAAACCTTGGATCAGATTTGCTCTCTAAAGTGATATAA 258
QY 2863 TTTTGCACTGTGTAGTCTCTGTGTATACATCTGGCTATTTAAACCAATGAGGAGTTTA 2922
DB TTTTGCACTGTGTAGTCTCTGTGTATACATCTGGCTATTTAAACCAATGAGGAGTTTA 198
QY 2923 ATCACAGAAACACAGATATGATCATATTTTGATACCTGCCATGTAAATCTACTC 2982
DB ATCACAGAAACACAGATATGATCATATTTTGATACCTGCCATGTAAATCTACTC 138
QY 2983 CTGAATAATAATGTGGTGCCATGACGGGTCTACGGTTTGTGGTAGTAACTTAATACCTT 3042
DB CTGAATAATAATGTGGTGCCATGACGGGTCTACGGTTTGTGGTAGTAACTTAATACCTT 78
QY 3043 AACCCCATGCTCAAAATCAAAATGATATATCTCTGAGACCCAGCAATACCATAG 3102
DB AACCCCATGCTCAAAATCAAAATGATATATCTCTGAGACCCAGCAATACCATAG 18
QY 3103 AATTACTAAAAA 3119
DB AATTACTAAAAA 1

RESULT 129

DA455896 558 bp mRNA linear EST 07-NOV-2005
LOCUS DA455896 CTONG3 Homo sapiens cDNA clone CTONG3000794 5', mRNA
DEFINITION DA455896
ACCESSION DA455896
VERSION DA455896.1 GI:81156902
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo

REFERENCE

1 (bases 1 to 558)
Kimura K., Watanabe S., Suzuki Y., Ota T., Nishikawa T.,
Yamashita R., Yamamoto J., Sekine M., Tsuritani K., Wakaguri H.,
Ishii S., Sugiyama T., Saito K., Isono Y., Irie R., Kishida N.,
Yoneyama T., Otsuka K., Kanda K., Yokoi T., Kondo H., Wagatsuma M.,
Murakawa K., Ishida S., Ishibashi T., Takahashi Fujii A.,
Tanabe T., Nagai K., Kikuchi H., Nakai K., Isegai T. and Sugano S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)

TITLE

JOURNAL

PUBLISHED 16344560
CONTACT: Takao Isegai
FLJ Project (HRI Team)
Helix Research Institute

2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction;
Helix Research Institute (HRI); 5'-end one pass sequencing; HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.

FEATURES

Location/Qualifiers
1..558
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CTONG3000794"
/tissue_type="tongue, tumor tissue"
/clone_lib="CTONG3"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 14.0%; Score 437; DB 9; Length 558;
Best Local Similarity 99.8%; Pred. No. 2.3e-224;
Matches 557; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 8 AAAGCTCCGAGGCGCAAGCGCTGCTACTGCGCCCGCTGCTTTAGTCCCGTTCGC 67
DB 1 AAAGCTCCGAGGCGCAAGCGCTGCTACTGCGCCCGCTGCTTTAGTCCCGTTCGC 60
QY 68 CGCTCTGGTGTGTCACCGCGCGCGCGGAGGAGCCACTGCAACAGGACCGGAGTGA 127
DB 61 CGCTCTGGTGTGTCACCGCGCGCGCGGAGGAGCCACTGCAACAGGACCGGAGTGA 120
QY 128 GCGCGCGCAGCATGAAGCGCGCGAGCGCCCTCATAGCGCAC-GTCCGAGCGTCCGG 186
DB 121 GCGCGCGCAGCATGAAGCGCGCGAGCGCCCTCATAGCGCACGCTCGGCGTCCGG 180
QY 187 CGGCGCGCGGAGGAGAAATGCAACATGCGCAGCAGCAATGGAACAGACAGCTGGCT 246
DB 181 CGGCGCGCGGAGGAGAAATGCAACATGCGCAGCAGCAATGGAACAGACAGCTGGCT 240
QY 247 GTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAAATCAGAGCTGCTAA 306
DB 241 GTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAAATCAGAGCTGCTAA 300
QY 307 TTGAGAGCTTTTATGTGTAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTCCGAT 366
DB 301 TTGAGAGCTTTTATGTGTAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTCCGAT 360
QY 367 ACCAGAAAATATCATGGCTACATGATGGCTTAAGGACACCATGATTCATGTTGTGAAG 426
DB 361 ACCAGAAAATATCATGGCTACATGATGGCTTAAGGACACCATGATTCATGTTGTGAAG 420
QY 427 AGGAATCATCCAGATGGACTCATTTCAGACAGAAATCTATTACCTGCCATGTCTGGTAG 486
DB 421 AGGAATCATCCAGATGGACTCATTTCAGACAGAAATCTATTACCTGCCATGTCTGGTAG 480
QY 487 AACAGAAAATACACTGTTTATTTCTGAAATTTCCAAATATCAATAGAGCAGCTC 546
DB 481 AACAGAAAATACACTGTTTATTTCTGAAATTTCCAAATATCAATAGAGCAGCTC 540
QY 547 TTAATGCTCTCTTGGAG 564
DB 541 TTAATGCTCTCTTGGAG 558

RESULT 130

CM427231 726 bp mRNA linear EST 16-MAY-2004
LOCUS CM427231
DEFINITION 17000455488465 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CM427231
VERSION CM427231.1 GI:47414825
KEYWORDS EST.

[illegible]

Db	313	CTGCATCTTTTGCACTACCTTCAAGAAAAACCTTGATCACGTAATTTGCTCTAAAGTG	254	Matches	430;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	2857	ATATATTTTGACCTGTGAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGA	2916										
Db	253	ATATAATTTTGACCTGTGAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGA	194										
Qy	2917	GTTTTAATCAACAGAAAAACACAGAATTGATCATCATATTTTGATACCTGCCATGTAACAT	2976										
Db	193	GGTTTAAATCAACAGAAAAACACAGAATTGATCATCATATTTTGATACCTGCCATGTAACAT	134										
Qy	2977	CTACTCTGAAATAAATGTTGGTGCATGCGAGGGGTCTACGGTTTGGTGTAGTAATCTAA	3036										
Db	133	CTACTCTGAAATAAATGTTGGTGCATGCGAGGGGTCTACGGTTTGGTGTAGTAATCTAA	74										
Qy	3037	TACCTTAAACCCACATGCTCAAAATCAATGATACATATTCCTGAGAGACCCAGCAATAC	3096										
Db	73	TACCTTAAACCCACATGCTCAAAATCAATGATACATATTCCTGAGAGACCCAGCAATAC	14										
Qy	3097	CATAAGAATTACT 3109											
Db	13	CATAAGAATTACT 1											
RESULT 132													
DB358826/c													
LOCUS													
DEFINITION	DB358826	HEMBB1 Homo sapiens cDNA clone HEMBB1000150 3', mRNA	EST 10-DEC-2005										
ACCESSION	DB358826												
VERSION	DB358826.1	GI:83486188											
KEYWORDS	EST.												
SOURCE	Homo sapiens (human)												
ORGANISM	Homo sapiens (human)												
REFERENCE													
AUTHORS	Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Teuritani,K., Wakesguri,H., Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuna,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A., Tanabe,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.												
TITLE	Diversification and Characterization of Putative Alternative Promoters of Human Genes												
JOURNAL	Genome Res. 16 (1), 55-65 (2006)												
PUBMED	16344560												
COMMENT	Contact: Takao Isogai FLJ Project (HRI Team) Helix Research Institute 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: flj-cdna@nifty.com HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute (HRI); cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and HRI.												
FEATURES													
Source	1. .514 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="HEMBB1000150" /tissue_type="whole embryo, mainly body" /dev_stage="embryo, 10 weeks" /clone_lib="HEMBB1" /note="Vector: pME18SFL3"												
ORIGIN													
Query Match	13.8%;	Score	430;	DB	9;	Length	514;						
Best Local Similarity	100.0%;	Pred.	No. 1.4e-220;										

RESULT 133

DA456165

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

CONTACT

TEL

FAX

EMAIL

PROJECT

LIBRARY

CONSTRUCTION

DEPARTMENT

OF

VIROLOGY

INSTITUTE

OF

MEDICAL

SCIENCE

UNIVERSITY

OF

TOKYO

AND

HRI.

LOCATION/QUALIFIERS

1. .514

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="HEMBB1000150"

/tissue_type="whole embryo, mainly body"

/dev_stage="embryo, 10 weeks"

/clone_lib="HEMBB1"

/note="Vector: pME18SFL3"

Query Match

Best Local Similarity

13.8%;

Score

430;

DB

9;

Length

514;

Pred.

No. 1.4e-220;

I (bases 1 to 546)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,

480 TGGTGGAGACAGAGAAAATACACTGTTTTTATTCTGAAATCCCAAAACTATCAATAGAGC 539

484 TGGTGGAGACAGAGAAAATACACTGTTTTTATTCTGAAATCCCAAAACTATCAATAGAGC 543

480 TGGTGGAGACAGAGAAAATACACTGTTTTTATTCTGAAATCCCAAAACTATCAATAGAGC 539

484 TGGTGGAGACAGAGAAAATACACTGTTTTTATTCTGAAATCCCAAAACTATCAATAGAGC 543

(3.5-4.4kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de Fatima Ronaldo, Ph.D. and M. Bento Soares, Ph.D. "

Query Match	13.3%	Score 416	DB 5	Length 474
Best Local Similarity	99.8%	Pred. No. 5,6e-213		
Matches 466	Conservative 0	Mismatches 1	Indels 0	Gaps 0
948	AGATGCCAGATCAGCTGGAGTCGCTACCTTTGGTTCTCCAGAAGAAATTCATAGATATTC	1007		
Qy				
Db	8 AGATGCCAGATCAGCTGGAGTCGCTACCTTTGGTTCTCCAGAAGAAATTCATAGATATTC	67		
Qy				
1008	TGGCTATTGGTGGTGTCGCAAAAGCTGAAACAACTCCCAAGTGGTGGTAAAAATTCCTAGAAT	1067		
Qy				
Db	68 TGGCTATTGGTGGTGTCGCAAAAGCTGAAACAACTCCCAAGTGGTGGTAAAAATTCCTAGAAT	127		
Qy				
1068	TCTATATGAAGAAAAATGATGAATCTGAGGTGGAATATTATCATGTTACATCCCTATGTT	1127		
Qy				
Db	128 TCTATATGAAGAAAAATGATGAATCTGAGGTGGAATATTATCATGTTACATCCCTATGTT	187		
Qy				
1128	GGAACAAAGAGGCGAGATTCATTCGGTTATCCTAAAAACGGTACAGCAAAATCCTAAAGT	1187		
Qy				
Db	188 GGAACAAAGAGGCGAGATTCATTCGGTTATCCTAAAAACGGTACAGCAAAATCCTAAAGT	247		
Qy				
1188	CACTTTTAAGATGTCAGAAATAATGATGTGATGCTGAAAGAGGATCATAGATGTCATAGA	1247		
Qy				
Db	248 CACTTTTAAGATGTCAGAAATAATGATGTGATGCTGAAAGAGGATCATAGATGTCATAGA	307		
Qy				
1248	TAAAGAACTAATTCACACTTTTGAGATTCCTATTTCGAAGGAGTTGAATAATATTGCCAGAGC	1307		
Qy				
Db	308 TAAAGAACTAATTCACACTTTTGAGATTCCTATTTCGAAGGAGTTGAATAATATTGCCAGAGC	367		
Qy				
1308	TGATGGACTCCTCAGAGGAAAAATATGTTTGGTCCATCTCTATAGATCGCTCCGACTCG	1367		
Qy				
Db	368 TGATGGACTCCTCAGAGGAAAAATATGTTTGGTCCATCTCTATAGATCGCTCCGACTCG	427		
Qy				
1368	CCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAAG	1414		
Qy				
Db	428 CCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAAG	474		
Qy				

RESULT	138
DAZ97f112	
LOCUS	570 bp mRNA linear EST 03-NOV-2005
DEFINITION	BRHIP2 Homo sapiens cDNA clone BRHIP2010255 5', mRNA sequence.

ACCESSION	DA297112
VERSION	DA297112.1
KEYWORDS	GI:79235254
SOURCE	EST.
ORGANISM	Homo sapiens (human)
PREFERENCE	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
	Hominidae; Homo.
	1 (bases 1 to 570)

AUTHORS
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, I.,
Yanase, K., Yamamoto, J., Sekine, M., Tsurtani, K., Wagauchi, H.,
Vanashita, R., Yamamoto, J., Sekine, M., Tsurtani, K., Wagauchi, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushiya, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Muraoka, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isozaki, T. and Sugano, S.
TITLE
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Vol. 46 (1) 55-65 (2006)

JOURNAL
PUBLISHED
COMMENT

Genome Res. 16:127-33 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986

209	Db	AGGAAAATGCAACATCGCAGCAGCAATAGGAAACAGACAGCTGGTGTGTGAGATATTGA	268
261	Qy	AACGTGCGGACTGTGAGGAGAAATATTGAATCAAGGATCGGCCCTAAATTTGAGAGCCTTTTGA	320
269	Db	AACGTGCGGACTGTGAGGAGAAATATTGAATCAAGGATCGGCCCTAAATTTGAGAGCCTTTTGA	328
321	Qy	TGTTTGAGCGGTATTCCTGGAGTCAGCTTTAAAAGCTGCTTGGCGATACACAGAAAATATCA	380
329	Db	TGTTTGAGCGGTATTCCTGGAGTCAGCTTTAAAAGCTGCTTGGCGATACACAGAAAATATCA	388
381	Qy	TGGGTACATGATGCTTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAAATGATCCAGA	440
389	Db	TGGGTACATGATGCTTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAAATGATCCAGA	448
441	Qy	TGGACCTCATTTACAGACAGAAATCTATTACCTTGCCATGTCTGGTCAGAACAGAGAAATAC	500
449	Db	TGGACCTCATTTACAGACAGAAATCTATTACCTTGCCATGTCTGGTCAGAACAGAGAAATAC	508
501	Qy	ACTGTTTATTTCTGAATATCCCAAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTG	560
509	Db	ACTGTTTATTTCTGAATATCCCAAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTG	568
561	Qy	GAAGCCTCTTTTGGATCTTTTTTTCAGGCACACTGGACTATGAAATGTATTCTCGAGAA	618
569	Db	GAAGCCTCTTTTGGATCTTTTTTTCAGGCACACTGGACTATGAAATGTATTCTCGAGAA	626

RESULT 137	CF137069	474 bp	mRNA	linear	EST 09-SEP-2003
LOCUS	UI-HP-ENO-amk-f-06-0-UI.r1	NIH_MGC_50	Homo sapiens	cDNA clone	
DEFINITION	IMAGE:3090562 5', mRNA sequence.				

ACCESSION	CF137069
VERSION	CF137069.1
KEYWORDS	GI:33252513
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
	Hominidae; Homo.
	1 (bases 1 to 474)
	Bonaldo,M.F., Lennon,G. and Soares,M.B.
	Normalization and subtraction: two approaches to facilitate gene
	discovery
	Genome Res. 6 (9): 791-806 (1996)

8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Louis Staudt
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/humanfl.html>
Seq primer: pYX-5.

FEATURES	source
Location/Qualifiers	
1. .474	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="IMAGE:3090562"	
/tissue_type="lymph"	
/cell_type="germinal center B cells"	
/cell_line="MGC85"	
/lab_host="DH10B (TTL)"	
/clone_lib="NIH_MGC 50"	
/note="vector: pT73D-PacI, Site 1: NotI, Site 2: Eco RI; Constructed from size fractionated cytoplasmic mRNA"	

Email: flj-cdna@nifty.com
 NEDO human cDNA project (New Energy and Industrial Technology
 Developmental Organization, Japan); cDNA library construction:
 Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
 Research Association for Biotechnology (RAB) and Biotechnology
 Center, National Institute of Technology and Evaluation; 3'-end one
 pass sequencing: RAB.

FEATURES
 source
 Location/Qualifiers
 1..570
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="BRH2010255"
 /tissue="hippocampus"
 /clone_lib="BRH1P2"
 /note="Vector: PME18SFL3"

ORIGIN
 Query Match 13.2%; Score 413; DB 9; Length 570;
 Best Local Similarity 100.0%; Pred. No. 2.4e-211;
 Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 146 GGCACAGCGCGCTCCATAGCGACGTCGGGACGGTCCGGGCGGGCGGGGGAAGGAA 205
 Db 158 GGCACAGCGCGCTCCATAGCGACGTCGGGACGGTCCGGGCGGGCGGGGGAAGAA 217
 QY 206 AATGCAACATGGCAGCAGCAATGGAACAGAAACAGACAGCTGGGTGTTGAGATATTTGAACTG 265
 Db 218 AATGCAACATGGCAGCAGCAATGGAACAGAAACAGACAGCTGGGTGTTGAGATATTTGAACTG 277
 QY 266 CGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCCTAAATTCGGACCTTTTATGTTG 325
 Db 278 CGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCCTAAATTCGGACCTTTTATGTTG 337
 QY 326 AGCGGTATTCCTGGAGTCAGCTTAAAGAGTCGTCGGATACCAAGAAATATCATGGCT 385
 Db 338 AGCGGTATTCCTGGAGTCAGCTTAAAGAGTCGTCGGATACCAAGAAATATCATGGCT 397
 QY 386 ACATGATGGCTAAGGCACACATGATTTTCATGTTGTGAGAGAAATGTCAGATGGAC 445
 Db 398 ACATGATGGCTAAGGCACACATGATTTTCATGTTGTGAGAGAAATGTCAGATGGAC 457
 QY 446 CTCATTACAGAGAAATCTATTACCTGCGATGTCGTTGAGACAGAGAAATACACTGT 505
 Db 458 CTCATTACAGAGAAATCTATTACCTGCGATGTCGTTGAGACAGAGAAATACACTGT 517
 QY 506 TTATTCTGAAATTCCTCCAAACTATCATAGACAGCAGCTCTTAATGCTCTCT 558
 Db 518 TTATTCTGAAATTCCTCCAAACTATCATAGACAGCAGCTCTTAATGCTCTCT 570

RESULT 139
 LOCUS DA551227 HCHON2 Homo sapiens cDNA clone HCHON2000347 5', mRNA
 DEFINITION DA551227
 ACCESSION DA551227.1 GI:80808232
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1. (bases 1 to 530)
 Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
 Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakisugi, H.,
 Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kuehida, N.,
 Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagaatsuma, M.,
 Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
 Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
 Diversification of Transcriptional Modulation: Large-scale
 Identification and Characterization of Putative Alternative

Promoters of Human Genes
 Genome Res. 16 (1), 55-65 (2006)
 16344560
 Contact: Takao Isogai
 FLJ Project (HRI Team)
 Helix Research Institute
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: flj-cdna@nifty.com
 NEDO human cDNA project (New Energy and Industrial Technology
 Developmental Organization, Japan); cDNA library construction:
 Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
 Research Association for Biotechnology (RAB) and Biotechnology
 Center, National Institute of Technology and Evaluation; 3'-end one
 pass sequencing: RAB.

FEATURES
 source
 Location/Qualifiers
 1..530
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="HCHON200347"
 /cell_type="chondrocytes (HC)"
 /clone_lib="HCHON2"
 /note="Vector: PME18SFL3; primary culture, chondrocytes"

ORIGIN
 Query Match 13.1%; Score 408; DB 9; Length 530;
 Best Local Similarity 99.8%; Pred. No. 1.2e-208;
 Matches 528; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 AAGTGTAAAGCCTCCGAGGCCAAGCCGCTCTACTCTCGCGCGCTCTTCTTAGTGC 60
 Db 2 AAGTGTAAAGCCTCCGAGGCCAAGCCGCTCTACTCTCGCGCGCTCTTCTTAGTGC 61
 QY 61 CGTTTCGCGCGCTGGGTGTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 Db 62 CGTTTCGCGCGCTGGGTGTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 121
 QY 121 GAGTGGAGGCG 179
 Db 122 GAGTGGAGGCG 181
 QY 180 GTCG 239
 Db 182 GTCG 241
 QY 240 GCTGGGTGTTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGATCG 299
 Db 242 GCTGGGTGTTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGATCG 301
 QY 300 GCTTAAATGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCT 359
 Db 302 GCTTAAATGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCT 361
 QY 360 TCCCGATACCAAGAAATATCATGCTACATGCTTAAGGCGCCACCATGATTTTCATGTT 419
 Db 362 TCCCGATACCAAGAAATATCATGCTACATGCTTAAGGCGCCACCATGATTTTCATGTT 421
 QY 420 TGTGAAGGAAATGATCCAGATGACCTCATTCAGACAGAAATCTATTACCTTCCCATGTC 479
 Db 422 TGTGAAGGAAATGATCCAGATGACCTCATTCAGACAGAAATCTATTACCTTCCCATGTC 481
 QY 480 TGTGAGAACACAGAGAAATACACTGTTTATTCCTGAAATTCCTGAAATTCCTGAAATTC 528
 Db 482 TGTGAGAACACAGAGAAATACACTGTTTATTCCTGAAATTCCTGAAATTCCTGAAATTC 530

RESULT 140
 LOCUS DA436183
 DEFINITION
 DA436183
 CTONG2 Homo sapiens cDNA clone CTONG2003086 5', mRNA
 linear
 EST 06-NOV-2005
 sequence.


```
ACCESSION DA436183
VERSION DA436183.1 GI:80867593
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 569)
AUTHORS Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaquri,H.,
Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
TITLE Identification of Transcriptional Modulation: Large-scale
Diversification and Characterization of Putative Alternative
Promoters of Human Genes
JOURNAL Genome Res. 16 (1), 55-65 (2006)
PUBMED 16344560
COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source           1..569
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CTONG2003086"
                     /tissue types="tongue, tumor tissue"
                     /clone_lib="CTONG2"
                     /note="Vector: pME18SFL3"
ORIGIN
Query Match      13.1%; Score 408; DB 9; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.2e-208;
Matches 408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 158 CTCATAGCGCACGTCGGGACGTCGCGGCGGGCGGGGGAAGGAAATGCAACATGG 217
DB 162 CTCATAGCGCACGTCGGGACGTCGCGGCGGGCGGGGGAAGGAAATGCAACATGG 221
QY 218 CAGCAGCAATGAAACAGACAGCTGGTGTTCAGATATTTGAACTGCGGACTGTGAGG 277
DB 222 CAGCAGCAATGAAACAGACAGCTGGTGTTCAGATATTTGAACTGCGGACTGTGAGG 281
QY 278 AGAATATTGAATCA CAGGATCGGCCTAAATTTGAGCGCTTTTATGTGTGAGCGGTATCT 337
DB 282 AGAATATTGAATCA CAGGATCGGCCTAAATTTGAGCGCTTTTATGTGTGAGCGGTATCT 341
QY 338 GGAGTCAGCTTAAAGAGCTGTCGCATACACAGAAATATCATGCTACATGATGGCTA 397
DB 342 GGAGTCAGCTTAAAGAGCTGTCGCATACACAGAAATATCATGCTACATGATGGCTA 401
QY 398 AGGCACCA CAGATGTTTATGTTGTGAAGAGGAATGATCCAGATGGACCTCAATT CAGACA 457
DB 402 AGGCACCA CAGATGTTTATGTTGTGAAGAGGAATGATCCAGATGGACCTCAATT CAGACA 461
QY 458 GAATCTATTACCTTGCCATGTCGTGAGAACACAGAGAAATACACTGTTTATTTCTGAAA 517
DB 462 GAATCTATTACCTTGCCATGTCGTGAGAACACAGAGAAATACACTGTTTATTTCTGAAA 521
QY 518 TTCCCAAAACTATCAATAGACGACGAGCTTTAATGCTCTTAAATGCTCTTTGGAAGC 565

Db 522 TTCCCAAAACTATCAATAGACGACGAGCTTTAATGCTCTTAAATGCTCTTTGGAAGC 569
|||||
RESULT 141
LOCUS DA946649
DEFINITION DA946649 SPLEN2 Homo sapiens cDNA clone SPLEN2022485 5', mRNA
sequence.
ACCESSION DA946649
VERSION DA946649.1 GI:83055100
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 590)
AUTHORS Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaquri,H.,
Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
TITLE Identification of Transcriptional Modulation: Large-scale
Diversification and Characterization of Putative Alternative
Promoters of Human Genes
JOURNAL Genome Res. 16 (1), 55-65 (2006)
PUBMED 16344560
COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
FEATURES             Location/Qualifiers
     source           1..590
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="SPLEN2022485"
                     /tissue type="spleen"
                     /clone_lib="SPLEN2"
                     /note="Vector: pME18SFL3"
ORIGIN
Query Match      12.9%; Score 404; DB 9; Length 590;
Best Local Similarity 100.0%; Pred. No. 1.8e-206;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 179 GGTCCGGGCGGGCGGGGGAAGGAAATGCAACATGCGCAGCAGCAATGGAACAGAAC 238
DB 187 GGTCCGGGCGGGCGGGGGAAGGAAATGCAACATGCGCAGCAGCAATGGAACAGAAC 246
QY 239 AGCTGGGTCTCAGATATTTGAACTGCGGACTGTGAGGAGAAATTTGAATTCACAGATC 298
DB 247 AGCTGGGTCTCAGATATTTGAACTGCGGACTGTGAGGAGAAATTTGAATTCACAGATC 306
QY 299 GGCCTAAATTTGAGCGCTTTTATGTGTGAGCGGTATTTCTTGAGTCAGCTTAAAAAGCTGC 358
DB 307 GGCCTAAATTTGAGCGCTTTTATGTGTGAGCGGTATTTCTTGAGTCAGCTTAAAAAGCTGC 366
QY 359 TTGCCATACCCAGAAATATCATGCTACATGCTGGCTAAGCCACCATGATTTTCATGT 418
DB 367 TTGCCATACCCAGAAATATCATGCTACATGCTGGCTAAGCCACCATGATTTTCATGT 426
```

```

QY 419 TTGTGAAGGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTTATTACCTTGCCATGT 478
Db 427 TTGTGAAGGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTTATTACCTTGCCATGT 486
QY 479 CTGGTGAGAACAGAGAAATACACTGTTTATTCTGAATTTCCCAAACTATCAATAGAG 538
Db 487 CTGGTGAGAACAGAGAAATACACTGTTTATTCTGAATTTCCCAAACTATCAATAGAG 546
QY 539 CAGCAGTCTTAATGCTCTCTCTGGAAGCCTCTTTTGATCTTTT 582
Db 547 CAGCAGTCTTAATGCTCTCTCTGGAAGCCTCTTTTGATCTTTT 590

RESULT 142
BI548490
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 835)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11661 row: f column: 01
High quality sequence stop: 795.
Location/Qualifiers
1..835
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5262336"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH MGC 95"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
(gtccag); Oligo dt primed using primer
5'-TTTATTTTATTTTATTTVN-3', size-selected for average
insert size 2.5 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this
is a NIH_MGC Library."

ORIGIN
Query Match
Best Local Similarity 12.9%; Score 404; DB 2; Length 835;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 201 AGGAAATGCAACATGCGACAGCAATGGAAACAGACAGCTGGGTGTGAGATATTGA 260
Db 229 AGGAAATGCAACATGCGACAGCAATGGAAACAGACAGCTGGGTGTGAGATATTGA 288
QY 261 AACTCGCGACTGTGAGAGAAATTTGAATCAGAGATCGCCCTAAATTTGGAGCCTTTT 320

```

```

Db 289 AACTCGCGACTGTGAGAGAAATTTGAATCAGAGATCGGCCTAAATTTGGAGCCTTTT 348
QY 321 TGTGAGCGGTATTCTCGAGTCAGCTTTAAAAAGCTGCTGCCGATACCAAGAAATATCA 380
Db 349 TGTGAGCGGTATTCTCGAGTCAGCTTTAAAAAGCTGCTGCCGATACCAAGAAATATCA 408
QY 381 TGGCTACATGATGGCTAAGGCCACACATGATTTTCATGTTTGTGAGAGGAAATGATCCAGA 440
Db 409 TGGCTACATGATGGCTAAGGCCACACATGATTTTCATGTTTGTGAGAGGAAATGATCCAGA 468
QY 441 TGGACCTCATTCAGACAGAACTTATTACCTTGCCATGCTCTGGTGAGAAAGAAATATC 500
Db 469 TGGACCTCATTCAGACAGAACTTATTACCTTGCCATGCTCTGGTGAGAAAGAAATATC 528
QY 501 ACTGTTTATTCTGAATTTCCCAAACTATCAATAGACAGCAGTCTTTAATGCTCTCTTTG 560
Db 529 ACTGTTTATTCTGAATTTCCCAAACTATCAATAGACAGCAGTCTTTAATGCTCTCTTTG 588
QY 561 GAAGCCTCTTTTGGATCTTTTCAGGCAACACTGGACTATGGAA 604
Db 589 GAAGCCTCTTTTGGATCTTTTCAGGCAACACTGGACTATGGAA 632

```

```

RESULT 143
AA769522/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 399)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 1011 Std Error: 0.00
Seq primer: -40ml3 fwd R from Amersham
High quality sequence stop: 394.
Location/Qualifiers
1..399
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1289774"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/clone_lib="NCI-CGAP GCBI"
/notes="Vector: p7T3D-Faci; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was prepared from human tonsillar cells
enriched for germinal center B cells by flow sorting
(CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr.
David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA
synthesis was primed with a Not I - oligo(dT) primer
[5'-TGTACCAATCTGAAGTGGAGCGCGCTCATTTTCTTTT-3',
1. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I

```

```

FEATURES
source

```

Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

FEATURES source

Location/Qualifiers
1. .599
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TRACH2015254"
/tissue_type="trachea"
/clone_lib="TRACH2"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 12.8%; Score 398; DB 9; Length 599;
Best Local Similarity 100.0%; Pred. No. 3.3e-203; Indels. 0; Gaps 0;
Matches 398; Conservative 0; Mismatches 0;
QY 170 CGTCGGGACGGTCCGGCGGGGGGGAAGGAAATGCAACATGGCAGCAGCAATGG 229
Db 167 CGTCGGGACGGTCCGGCGGGGGGGAAGGAAATGCAACATGGCAGCAGCAATGG 226
QY 230 AAACAGAACACGCTGGGTGTTGAGATATTTCAAACTGCGGACTGTGAGGAGATATTTGAAT 289
Db 227 AAACAGAACACGCTGGGTGTTGAGATATTTCAAACTGCGGACTGTGAGGAGATATTTGAAT 286
QY 290 CACAGGATCGGCTTAAATTTGAGGCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTA 349
Db 287 CACAGGATCGGCTTAAATTTGAGGCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTA 346
QY 350 AAAAGCTGTTGCCGATACCAAGAAATATCATGCTACATGATGGCTTAAGGACCCACATG 409
Db 347 AAAAGCTGTTGCCGATACCAAGAAATATCATGCTACATGATGGCTTAAGGACCCACATG 406
QY 410 ATTTTCATGTTTCTGAAGAGGAAATGATCCAGATGGACCTCATTCAGACAGAAATTTATACC 469
Db 407 ATTTTCATGTTTCTGAAGAGGAAATGATCCAGATGGACCTCATTCAGACAGAAATTTATACC 466
QY 470 TTGCCATGTCGTGGTGAAGAACAGAGAAATACACGTGTTTATTTCTGAAATTTCCCAAACTA 529
Db 467 TTGCCATGTCGTGGTGAAGAACAGAGAAATACACGTGTTTATTTCTGAAATTTCCCAAACTA 526
QY 530 TCAATAGACGACGCTTAAATGCTCTCTCTGGAAGCCT 567
Db 527 TCAATAGACGACGCTTAAATGCTCTCTCTGGAAGCCT 564

RESULT 145

CN427234 738 bp mRNA linear EST 16-MAY-2004
LOCUS 17000600021819 GRN_PRENUE Homo sapiens cDNA 5', mRNA sequence.
DEFINITION
ACCESSION CN427234
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Lebkyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 738)
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkyota, J and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)

REFERENCE

AUTHORS
TITLE
JOURNAL
PUBMED
COMMENT
Contact: Brandenberger R
Regenerative Medicine
Genon Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760

and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Ronaldo."

ORIGIN

Query Match 12.8%; Score 399; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 9.2e-204; Indels 0; Gaps 0;
Matches 399; Conservative 0; Mismatches 0;
QY 2720 GAAAGCCATATGATTACAGATCTATCTCCAGAGAGACACAGCATAAGAGTTCTCTGAAT 2779
Db 399 GAAAGCCATATGATTACAGATCTATCTCCAGAGAGACACAGCATAAGAGTTCTCTGAAT 340
QY 2780 CGGAGAACATTATGAACTGCATCTTTTGCACCTCAAGAAACCTTGATCAGTA 2839
Db 339 CGGAGAACATTATGAACTGCATCTTTTGCACCTCAAGAAACCTTGATCAGTA 280
QY 2840 TTGCTGCTTAAAGTGATATAATTTTGAACCTGTGAGAACCTCTCTGGTATACACTGGCT 2899
Db 279 TTGCTGCTTAAAGTGATATAATTTTGAACCTGTGAGAACCTCTCTGGTATACACTGGCT 220
QY 2900 ATTTAAACCAATGAGGAGGTTTAAATCAACAGAAACACAGAAATTCATCACATTTTGA 2959
Db 219 ATTTAAACCAATGAGGAGGTTTAAATCAACAGAAACACAGAAATTCATCACATTTTGA 160
QY 2960 TACCTGCCATGTAACATCTACTCTCGAAATAAATGTGGTCCATGCGAGGGTCTACGGT 3019
Db 159 TACCTGCCATGTAACATCTACTCTCGAAATAAATGTGGTCCATGCGAGGGTCTACGGT 100
QY 3020 TTGTGTGAGTAATCTAATACCTTAACCCCAATGCTCAAAATCAAAATGATATTCCT 3079
Db 99 TTGTGTGAGTAATCTAATACCTTAACCCCAATGCTCAAAATCAAAATGATATTCCT 40
QY 3080 GAGAGACCCAGCAATACCATAGAAATTAACAAAAA 3118
Db 39 GAGAGACCCAGCAATACCATAGAAATTAACAAAAA 1

RESULT 144

DB204093 599 bp mRNA linear EST 04-DEC-2005
LOCUS 16344560
DEFINITION TRACH2 Homo sapiens cDNA clone TRACH2015254 5', mRNA sequence.

ACCESSION DB204093
VERSION DB204093.1 GI:83245645
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 599)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi, Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification and Characterization of Putative Alternative Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)

TITLE

JOURNAL
PUBMED
COMMENT
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,

Email: rbrandenberger@geron.com
Insert Length: 738 Std Error: 0.00.
Location/Qualifiers
1. .738
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/issue_type="embryonic stem cell, retinoic acid and
mitogen-treated HES cell line H7"
/clone_lib="GRN PRENEU"
/note="oligo at primed, full-length enriched cDNA library
from HES cell line H7, (p29) maintained in feeder-free
conditions. Embryoid bodies were generated in the presence
of all-trans retinoic acid and mitogens."

FEATURES
source

ORIGIN

Query Match 12.7%; Score 395; DB 8; Length 738;
Best Local Similarity 100.0%; Pred. No. 1.4e-201;
Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1938 GAAGATCCACCTGTGTGTCCTTTTAAAGTATCAAGTCTTCAAGATGACCAACTTG 1997
Db 9 GAAGATCCACCTGTGTGTCCTTTTAAAGTATCAAGTCTTCAAGATGACCAACTTG 68
QY 1998 CAAACAAAGGAATTTGGGCCACCATTTGGATTTCAGCAGGTCTCTTCCGACTATAC 2057
Db 69 CAAACAAAGGAATTTGGGCCACCATTTGGATTTCAGCAGGTCTCTTCCGACTATAC 128
QY 2058 TCCTCCAGAAATTTCTCTTTGAAAGTACTACTGGATTTACATTTGATGGATGCTCTA 2117
Db 129 TCCTCCAGAAATTTCTCTTTGAAAGTACTACTGGATTTACATTTGATGGATGCTCTA 188
QY 2118 CAAGCTCATGATCTACAGCTGGAAAGAAATATCTACTGTCTGTTCATATATGGTGG 2177
Db 189 CAAGCTCATGATCTACAGCTGGAAAGAAATATCTACTGTCTGTTCATATATGGTGG 248
QY 2178 TCCTCAGGTGAGTTGGTGAATATCGTTTAAAGAGTCAAGTATTTCCGCTGAATAC 2237
Db 249 TCCTCAGGTGAGTTGGTGAATATCGTTTAAAGAGTCAAGTATTTCCGCTGAATAC 308
QY 2238 CTTAGCTCTCTAGTTATGTTGTAGTATAGACACAGGATCTGTCCAGG 2297
Db 309 CTTAGCTCTCTAGTTATGTTGTAGTATAGACACAGGATCTGTCCAGG 368
QY 2298 GCTTAAATTTGAAGCGCCTTTAAATATATATGG 2332
Db 369 GCTTAAATTTGAAGCGCCTTTAAATATATATATGG 403

RESULT 146

AA278625 493 bp mRNA linear EST 15-AUG-1997
LOCUS z878911.1 NCI CGAP GCBI Homo sapiens cDNA clone IMAGE:703652 5',
DEFINITION similar to SW:DPP4_RAT P14740 DIPEPTIDYL PEPTIDASE IV ;, mRNA
sequence.
AA278625
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 493)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TUMOR Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 817 Std Error: 0.00

Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 493.
Location/Qualifiers
1. .493
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:703652"
/issue_type="germlinal center B cell"
/lab_host="DH10B"
/clone_lib="NCI CGAP GCBI"
/notes="Vector: pVT3B-PacI; Site 1: Not I; Site 2: Eco RI;
1st strand cDNA was prepared from human tonsillar cells
enriched for germinal center B cells by flow sorting
(CD20+, IGD-), provided by Dr. Louis M. Staudt (NCI), Dr.
David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA
synthesis was primed with a Not I oligo (GT TTTT-3',
[5'-NGTTACCAATCTGAAGTGGGAGGCGCTCATTTTCTTTT-3',
Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

FEATURES
source

ORIGIN

Query Match 12.6%; Score 393; DB 1; Length 493;
Best Local Similarity 100.0%; Pred. No. 1.7e-200;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2333 GTCAAATGAAATTCAGCATCAGGTGGAAGCACTCCAATATCTAGCTTCTCGATATGATT 2392
Db 1 GTCAAATGAAATTCAGCATCAGGTGGAAGCACTCCAATATCTAGCTTCTCGATATGATT 60
QY 2393 TCATTGACTTATAGATCGTGTGGGCATCCACGGCTGGTCTATGGAGGATACCTCTCCCTGA 2452
Db 61 TCATTGACTTATAGATCGTGTGGGCATCCACGGCTGGTCTATGGAGGATACCTCTCCCTGA 120
QY 2453 TGGCAATTAATGACAGGTTCAGATATCTTCAGGGTTGCTATTGTGGGGCCCCAGTCACTC 2512
Db 121 TGGCAATTAATGACAGGTTCAGATATCTTCAGGGTTGCTATTGTGGGGCCCCAGTCACTC 180
QY 2513 TGTGATCTTCTATGATACAGGATACACGGAACTTATATGGGTCAACCTGACCAAGATG 2572
Db 181 TGTGATCTTCTATGATACAGGATACACGGAACTTATATGGGTCAACCTGACCAAGATG 240
QY 2573 AACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGACCAA 2632
Db 241 AACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGACCAA 300
QY 2633 ATCGTTTACTGCTCTTACATGTTTCTCTGATGAGATGTCCTTTCACATACCACTA 2692
Db 301 ATCGTTTACTGCTCTTACATGTTTCTCTGATGAGATGTCCTTTCACATACCACTA 360
QY 2693 TATTACTGAGTTTTTTAGTCAGGGCTGGAAGC 2725
Db 361 TATTACTGAGTTTTTTAGTCAGGGCTGGAAGC 393

RESULT 147

AA312210/c
LOCUS AA312210.1 Jurkat T-cells VI Homo sapiens cDNA 5' end, mRNA
DEFINITION EST182961
sequence.
ACCESSION AA312210
VERSION AA312210.1 GI:1964610
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 392)
REFERENCE Adams M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,
AUTHORS

Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitchugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudak, D.M., Shirley, R., Small, K.V., Spryger, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Damke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weisener, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.		Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence Nature 377 (6547 Suppl.), 3-174 (1995) 7566098	
Other ESTs: THC184894 Contact: Kerlavage, AR Bioinformatics The Institute for Genomic Research 9712 Medical Center Drive, Rockville, MD 20850 USA Tel: 3018699056 Fax: 3018699423 Email: arkerlav@tigr.org		For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (http://www.tigr.org/cdb/hgi/hgi.html) Seq primer: M13 Reverse. Location/Qualifiers 1. .392 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="tacc (inhost):159455" /db_xref="taxon:9606" /cell_type="T-lymphocyte" /clone_lib="Jurkat T-cells VI" /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI"	
TITLE JOURNAL PUBMED COMMENT		source	
FEATURES		Location/Qualifiers	
source		1. .392 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="tacc (inhost):159455" /db_xref="taxon:9606" /cell_type="T-lymphocyte" /clone_lib="Jurkat T-cells VI" /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI"	
ORIGIN		Query Match 12.6%; Score 392; DB 1; Length 392; Best Local Similarity 100.0%; Pred. No. 5.7e-200; Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	2687	CCAGTATATTACTGAGTCTTTTGTAGTGGGCTGGAAGCCATATGATTTACAGATCTATC	2746
DB	392	CCAGTATATTACTGAGTCTTTTGTAGTGGGCTGGAAGCCATATGATTTACAGATCTATC	333
QY	2747	CTCAGGAGAGACACAGCATAGAGTCTCTGAATCGGAGACATTTATGAATGCACTCTT	2806
DB	332	CTCAGGAGAGACACAGCATAGAGTCTCTGAATCGGAGACATTTATGAATGCACTCTT	273
QY	2807	TGCATCTCTTCAAGAAACCTTGATCAGTATTCCTGCTCTTAAAGTGATATATTT	2866
DB	272	TGCATCTCTTCAAGAAACCTTGATCAGTATTCCTGCTCTTAAAGTGATATATTT	213
QY	2867	GACCTGTGTAGAACTCTGTGTATACACTGGCTATTTAACCAATGAGGAGTAAATCA	2926
DB	212	GACCTGTGTAGAACTCTGTGTATACACTGGCTATTTAACCAATGAGGAGTAAATCA	153
QY	2927	ACAGAAAAACACAGAAATGATCATCACATTTTGTATCTGCGCATGTAACATCTACTCTGA	2986
DB	152	ACAGAAAAACACAGAAATGATCATCACATTTTGTATCTGCGCATGTAACATCTACTCTGA	93
QY	2987	AAATAAATGTGGTCCATGACAGGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAACC	3046
DB	92	AAATAAATGTGGTCCATGACAGGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAACC	33
QY	3047	CCACATGCTCAAAATCAAAATGATACATATTCC	3078
32		CCACATGCTCAAAATCAAAATGATACATATTCC	
RESULT 148		DB058900	
LOCUS		DEFINITION	
DB058900		552 bp mRNA linear EST 10-DEC-2005	
TEST12		Homo sapiens cDNA clone TEST12051728 5', mRNA	
sequence.		DB058900	
VERSION		DB058900	
KEYWORDS		EST. DB058900.1 GI:83509547	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		1 (bases 1 to 552) Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaquri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi, Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.	
TITLE		Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes	
JOURNAL		Genome Res. 16 (1), 55-65 (2006)	
PUBMED		16344560	
COMMENT		Contact: Takao Isogai FLJ Project (HRI Team) Helix Research Institute 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: flj-cdna@nifty.com NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction; Helix Research Institute (HRI); 5'-end one pass sequencing; HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing; RAB.	
FEATURES		Location/Qualifiers	
source		1. .552 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="tacc (inhost):9606" /clone="TEST12051728" /tissue_type="testis" /clone_lib="TEST12" /note="Vector: pME18SFL3"	
ORIGIN		Query Match 12.5%; Score 391; DB 9; Length 552; Best Local Similarity 100.0%; Pred. No. 2e-199; Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	832	GCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTA	891
DB	162	GCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTA	221
QY	892	ACCAGAGAGAAGAGAGACTCACTTATGTGCACATAGCTAGCCACATGGAAGAGAT	951
DB	222	ACCAGAGAGAAGAGAGACTCACTTATGTGCACATAGCTAGCCACATGGAAGAGAT	281
QY	952	GCCAGATCAGCTGGAGTCGCTACCTTGTCTCCAGAGAAATTTGATAGATATCTTGGC	1011
DB	282	GCCAGATCAGCTGGAGTCGCTACCTTGTCTCCAGAGAAATTTGATAGATATCTTGGC	341
QY	1012	TATTGTGGTGTCCAAAAGCTGAAACAACTCCAGTGGTGTGTAATAATCTTTAGAAATCTA	1071
DB	342	TATTGTGGTGTCCAAAAGCTGAAACAACTCCAGTGGTGTGTAATAATCTTTAGAAATCTA	401

```

QY 1072 TATGAAGAAATGATGAATCTGAGGTGGAATATTATTCATCTACATCCCTATGTTGGA 1131
DB |||||
QY 402 TATGAAGAAATGATGAATCTGAGGTGGAATATTATTCATCTACATCCCTATGTTGGA 461
DB |||||
QY 1132 ACAAGGAGGCGAGATTCAATCCGTTATCTTAAACAGGTACAGCAATCTTAAGTCACT 1191
DB |||||
QY 462 ACAAGGAGGCGAGATTCAATCCGTTATCTTAAACAGGTACAGCAATCTTAAGTCACT 521
DB |||||
QY 1192 TTTAAGATGTCAGAAATTAATGATGATGCTG 1222
DB |||||
QY 522 TTTAAGATGTCAGAAATTAATGATGATGCTG 552
DB |||||

RESULT 149
BG719202 798 bp mRNA linear EST 08-MAY-2001
LOCUS 603690257F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4822550 5',
DEFINITION mRNA sequence.
ACCESSION BG719202
VERSION BG719202.1 GI:13998389
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 798)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10731 row: a column: 15
High quality sequence stop: 796.

FEATURES
source
1..798
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4822550"
/lab_host="NH108"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescriptR (modified
pUCtag); Site_1: BamHI; Site_2: SalI-XhoI
5'-TTTTTTTTTTTNN-3', size-selected for average
insert size 2.2 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci,
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN
Query Match 12.4%; Score 388; DB 2; Length 798;
Best Local Similarity 99.8%; Pred. No. 8.9e-198;
Matches 438; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 201 AGGAAATGCAACATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTTGAGATATTGA 260
DB |||||
QY 359 AGGAAATGCAACATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTTGAGATATTGA 418
DB |||||
QY 261 AACTCGGAGTGTGAGGAGAAATTTGAATCAGAGGATCGGCCTTAATTCGAGCCCTTTTGA 320
DB |||||
QY 419 AACTCGGAGTGTGAGGAGAAATTTGAATCAGAGGATCGGCCTTAATTCGAGCCCTTTTGA 478
DB |||||

```

```

QY 321 TGTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTTCCGATACCGAATAATCA 380
DB |||||
QY 479 TGTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTTCCGATACCGAATAATCA 538
DB |||||
QY 381 TGGCTACATGATGGCTTAAGGCACCAATGATTTTCATGTTTGTGAAGGAGGATGATCCAGA 440
DB |||||
QY 539 TGGCTACATGATGGCTTAAGGCACCAATGATTTTCATGTTTGTGAAGGAGGATGATCCAGA 598
DB |||||
QY 441 TGGACCTCATTTCAGACAGCAATCTATACCTTGCATCTCTGGTGTAGAACAGAGAAATAC 500
DB |||||
QY 599 TGGACCTCATTTCAGACAGCAATCTATACCTTGCATCTCTGGTGTAGAACAGAGAAATAC 658
DB |||||
QY 501 ACTGTTTATTCCTGAAATCCCAAACTATCATATAGAGCAGCAGTCTTAATGCTCTCTTG 560
DB |||||
QY 659 ACTGTTTATTCCTGAAATCCCAAACTATCATATAGAGCAGCAGTCTTAATGCTCTCTTG 718
DB |||||
QY 561 GAAGCCTCTTTTGGATCTTTTTCAGGCAACATGAGACTATGGAATGATTTCTCGAGAAGA 620
DB |||||
QY 719 GAAGCCTCTTTTGGATCTTTTTCAGGCAACATGAGACTATGGAATGATTTCTCGAGAAGA 778
DB |||||
QY 621 AGAATCTATTAAGAGAAAGA 639
DB |||||
QY 779 AGAATCTATTAAGAGAAAGA 797
DB |||||

RESULT 150
DA184650 555 bp mRNA linear EST 03-NOV-2005
LOCUS DA184650 BRAMY3 Homo sapiens cDNA clone BRAMY3003323 5', mRNA
DEFINITION sequence.
ACCESSION DA184650
VERSION DA184650.1 GI:79178555
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 555)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakisugi, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Megatsuma, M.,
Murakawa, K., Iehida, S., Ishibashi, T., Takahashi, R., Tanase, T.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Tsogai, J., and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Katsuragi, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation, 3'-end one
pass sequencing: RAB.
Location/Qualifiers
1..555
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRAMY3003323"
/tissue_type="amygdala"
/clone_lib="BRAMY3"
/note="Vector: pME18SFL3"

```

Thu Jun 22 09:04:26 2006

ORIGIN

Query Match 12.3%; Score 384; DB 9; Length 555;
Best Local Similarity 99.8%; Fred. No. 1.3e-195; Indels 0; Gaps 0;
Matches 434; Conservative 0; Mismatches 1;

Qy	106	CTGCAACCAAGGACCGGAGTGGAGCGCGCAGCATGAAGCGGCGCAGGCCCGCTCCATAG	165
Db	102	CTGCAACCAAGGACCGGAGTGGAGCGCGCAGCATGAAGCGGCGCAGGCCCGCTCCATAG	161
Qy	166	CGCACGTCCGGACCGTCCGGCGCGGCGCGGCGGGAAGAAATGCAACATGGCAGCAGCA	225
Db	162	CGCACGTCCGGACCGTCCGGCGCGGCGCGGCGGGAAGAAATGCAACATGGCAGCAGCA	221
Qy	226	ATGGAACACAGACAGCTGGGTCTTGAGATATTTGAAACTGCGGACTGTGAGGAGAATATT	285
Db	222	ATGGAACACAGACAGCTGGGTCTTGAGATATTTGAAACTGCGGACTGTGAGGAGAATATT	281
Qy	286	GAATCACAGGATCGGCCTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAG	345
Db	282	GAATCACAGGATCGGCCTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAG	341
Qy	346	CTTAAAAAGCTGCTTCCGATACCAAGAAATATCATGGCTACATGATGGCTAAGGCACCA	405
Db	342	CTTAAAAAGCTGCTTCCGATACCAAGAAATATCATGGCTACATGATGGCTAAGGCACCA	401
Qy	406	CATGATTTCAATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTAT	465
Db	402	CATGATTTCAATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTAT	461
Qy	466	TACCTTGCCCATGCTCGTGAGAACAGAGAAATACACTGTTTATTCGAAATTCCTCCAAA	525
Db	462	TACCTTGCCCATGCTCGTGAGAACAGAGAAATACACTGTTTATTCGAAATTCCTCCAAA	521
Qy	526	ACTATCAATAGAGCA	540
Db	522	ACTATCAATAGAGCA	536

Search completed: June 22, 2006, 08:05:45
Job time : 14083 secs

THIS PAGE BLANK (USPTO)

Thu Jun 22 09:04:25 2006

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 22, 2006, 03:55:04 ; Search time 17271 Seconds
(without alignments)
11552.079 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 3120
Sequence: 1 aagtgttaagctccgagg.....agaattactataaaaaaaa 3120

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 6366136 seqs, 31973710525 residues

Word size : 1
Total number of hits satisfying chosen parameters: 12730834

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 150 summaries

Database : GenEmbl.*

- 1: gb env.*
- 2: gb pat.*
- 3: gb pi.*
- 4: gb pi.*
- 5: gb pr.*
- 6: gb ro.*
- 7: gb sts.*
- 8: gb sv.*
- 9: gb un.*
- 10: gb vi.*
- 11: gb ov.*
- 12: gb htg.*
- 13: gb in.*
- 14: gb om.*
- 15: gb ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3120	100.0	3120	5	AR651452 Sequence
2	3120	100.0	3127	5	AF221634 Homo sapi
3	3069	98.4	3127	2	CS196102 Sequence
4	3003	96.2	3106	2	AX342633 Sequence
5	2957	94.8	3143	2	AX354793 Sequence
6	2889	92.6	4829	2	AR631283 Sequence
7	2889	92.6	4829	2	AX608735 Sequence
8	2756	88.3	3030	5	AX354202 Homo sapi
9	2547	81.6	2649	2	AX354795 Sequence
10	2495	80.0	2656	5	AY172659 Sequence
11	2495	80.0	2671	2	AR631278 Sequence
12	2495	80.0	2671	2	AX608725 Sequence
13	2315	74.2	2797	2	AR448400 Sequence
14	2128	68.2	4523	2	AR631281 Sequence
15	2128	68.2	4523	2	AX608731 Sequence
16	2128	68.2	4676	2	AR631287 Sequence
17	2128	68.2	4676	2	AX608743 Sequence
18	2116	67.8	2842	2	AX405770 Sequence

19	2081	66.7	3125	5	BC030688
20	1984	63.6	4685	2	AR631288
21	1984	63.6	4685	2	AX608745
22	1963	62.9	2778	5	AK000290
23	1450	46.5	2161	2	BD157001
24	1450	46.5	2161	2	AX878058
25	1450	46.5	2161	5	AK027826
26	1377	44.1	2668	2	AX405771
27	1191	38.2	4309	2	AR631284
28	1191	38.2	4309	2	AX608737
29	1164	37.3	2510	2	AX338497
30	1023	32.8	1669	5	AF221636
31	1023	32.8	1669	5	AF221637
32	878	28.1	2098	2	CO721272
33	877	28.1	1356	2	AR631282
34	877	28.1	1356	2	AX608733
35	790	25.3	1197	5	AR651453
36	790	25.3	1197	5	AF221635
37	789	25.3	1083	2	AR651455
38	789	25.3	1083	5	AF221637
39	760	24.4	832	2	AR631286
40	760	24.4	832	2	AX608741
41	739	23.7	735	2	AX083130
42	708	22.7	735	2	AX524942
43	700	22.4	753	2	AX524935
44	672	21.5	1002	2	AX524939
45	501	16.1	631	2	BD149852
46	501	16.1	631	2	AX869790
47	472	15.1	620	2	AR631285
48	472	15.1	620	2	AX608739
49	383	12.3	198295	5	AC015129
50	339	10.9	147760	5	AC011846
51	209	9.9	823	2	AR263393
52	205	6.6	587	2	BD154822
53	205	6.6	587	2	AX874760
54	188	6.0	308	2	AX970984
55	188	6.0	308	2	BD109703
56	188	6.0	308	2	AR414150
57	167	5.4	168	2	AX522584
58	156	5.0	176282	12	AC068507
59	82	2.6	89	2	BD036103
60	82	2.6	89	2	AR736364
61	82	2.6	89	2	AX900570
62	82	2.6	477	7	BV448370
63	81	2.6	222034	12	AC157219
64	81	2.0	224227	12	AC106509
65	56	1.8	4799	6	BC043124
66	56	1.8	4799	6	BC059222
67	54	1.7	305843	12	AC144143
68	48	1.5	305843	12	AC144143
69	47	1.5	186144	6	AC161443
70	46	1.5	60	2	CQ545450
71	44	1.4	197073	12	AC171824
72	42	1.3	116580	6	AC112162
73	42	1.3	155733	6	AC161366
74	42	1.3	162528	6	AC122463
75	34	1.1	612	2	AR372392
76	33	1.1	72149	12	AC122776
77	32	1.0	1449	11	CR353491
78	29	0.9	194188	6	AC163387
79	26	0.8	2200	11	BC108576
80	25	0.8	25	2	CS196106
81	25	0.8	51	2	CQ002331
82	24	0.8	9154	2	CQ572930
83	24	0.8	40388	2	CQ608996
84	24	0.8	79001	12	AC015386
85	24	0.8	102653	4	AC090436
86	24	0.8	127686	4	AC090434
87	24	0.8	172372	13	AC010110
88	24	0.8	298235	12	AC127924
89	24	0.8	301691	13	AE003475
90	23	0.7	23	2	CS196104
91	23	0.7	23	2	AR631300

Qy 841 GACTGGATTGCTTTTATATACATAGCAACGATATTTGGATATCTAAATCGTAAACAGAGAA 900
Db 841 GACTGGATTGCTTTTATATACATAGCAACGATATTTGGATATCTAAATCGTAAACAGAGAA 900
Qy 901 GAAAGGAGACTCACATTATGTGCACAAATGAGCTAGCCACATCGGAAGAGATGCCATCA 960
Db 901 GAAAGGAGACTCACATTATGTGCACAAATGAGCTAGCCACATCGGAAGAGATGCCATCA 960
Qy 961 GCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATCTGGCTATTTGGTG 1020
Db 961 GCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATCTGGCTATTTGGTG 1020
Qy 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA 1080
Db 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA 1080
Qy 1081 AATGATGAATCTGAGGTGGAATTAATTCATGTATCATCTCCCTATGTTGGAAAAACAGGAGG 1140
Db 1081 AATGATGAATCTGAGGTGGAATTAATTCATGTATCATCTCCCTATGTTGGAAAAACAGGAGG 1140
Qy 1141 GCAGATTCATTCGGTTATCTCTAAACAGGTAACAGTAACAGTAACAGTAACAGTAACAGTA 1200
Db 1141 GCAGATTCATTCGGTTATCTCTAAACAGGTAACAGTAACAGTAACAGTAACAGTAACAGTA 1200
Qy 1201 TCAGAAATTAATGATGATGCTGGAAGGAAGATCATAGATGTCATAGATAAGGAACCTAAT 1260
Db 1201 TCAGAAATTAATGATGATGCTGGAAGGAAGATCATAGATGTCATAGATAAGGAACCTAAT 1260
Qy 1261 CAACTTTTGAAGATCTTATTTGAAGGAGTTGAAATATATGTCAGAGCTGGATGCACTCCT 1320
Db 1261 CAACTTTTGAAGATCTTATTTGAAGGAGTTGAAATATATGTCAGAGCTGGATGCACTCCT 1320
Qy 1321 GAGGGAATATGCTTGTGCTCAATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTG 1380
Db 1321 GAGGGAATATGCTTGTGCTCAATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTG 1380
Qy 1381 TTGATCTCACCTGAAATATTTATCCAGTGAAGAATGATGTTATGGAAGGAGGAGACTC 1440
Db 1381 TTGATCTCACCTGAAATATTTATCCAGTGAAGAATGATGTTATGGAAGGAGGAGACTC 1440
Qy 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAATATCTATGAAGAAAAACACAGACATC 1500
Db 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAATATCTATGAAGAAAAACACAGACATC 1500
Qy 1501 TGGATAATATCCATGACATCTTTTCAATGTTTTTCCCAAGTACGAAAGGAAATTTGAG 1560
Db 1501 TGGATAATATCCATGACATCTTTTCAATGTTTTTCCCAAGTACGAAAGGAAATTTGAG 1560
Qy 1561 TTTATTTTTGCTCTGAAATGCAAAACAGGTTTTCCGTCAATTTATACAAAATTAATCTATT 1620
Db 1561 TTTATTTTTGCTCTGAAATGCAAAACAGGTTTTCCGTCAATTTATACAAAATTAATCTATT 1620
Qy 1621 TTTAAAGGAAGCAAAATAAAACGATCCAGTGGTGGCTGCTGCTCCAGTGAATTTCAAG 1680
Db 1621 TTTAAAGGAAGCAAAATAAAACGATCCAGTGGTGGCTGCTGCTCCAGTGAATTTCAAG 1680
Qy 1681 TGTCTCTATCAAAAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTGGCCGGCAT 1740
Db 1681 TGTCTCTATCAAAAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTGGCCGGCAT 1740
Qy 1741 GGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACAAAGAC 1800
Db 1741 GGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACAAAGAC 1800
Qy 1801 TCCCTTTTAGAGCATCACTGTATCGTATGATGTTAATTAATCTGGAGAGGTGCAAGG 1860
Db 1801 TCCCTTTTAGAGCATCACTGTATCGTATGATGTTAATTAATCTGGAGAGGTGCAAGG 1860
Qy 1861 CTGACTGACCGGGCTACTCAATTTCTGCTGCAATCAGTCAGGACTGTGACTCTTTTATA 1920
Db 1861 CTGACTGACCGGGCTACTCAATTTCTGCTGCAATCAGTCAGGACTGTGACTCTTTTATA 1920
Qy 1921 AGTAAAGTATAGTAAACAGAGAAATCCACATGTGTGTCCTTTTACAAGCTATCAAGTCCT 1980

Db 1921 AGTAAAGTATAGTAAACAGAGAAATCCACATGTGTGTCCTTTTACAAGCTATCAAGTCCT 1980
Qy 1981 GAAGATGACCAACTTGGCAAAACAAAGGAAATTTTGGGCCACCATTTTGGATTTCAAGCAGGT 2040
Db 1981 GAAGATGACCAACTTGGCAAAACAAAGGAAATTTTGGGCCACCATTTTGGATTTCAAGCAGGT 2040
Qy 2041 CCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTAATCTACTGGATTTTACA 2100
Db 2041 CCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTAATCTACTGGATTTTACA 2100
Qy 2101 TTGTATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTG 2160
Db 2101 TTGTATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTG 2160
Qy 2161 CTGTTTCAATATATGCTGCTCTCAGGTCAGTTGGTGAATTAATCGGTTTAAAGGAGTCAAG 2220
Db 2161 CTGTTTCAATATATGCTGCTCTCAGGTCAGTTGGTGAATTAATCGGTTTAAAGGAGTCAAG 2220
Qy 2221 TATTTCCGCTTGAATACCTCTAGCTCTCTAGGTTATGTTGGTGTAGTATGATAGACAACAGG 2280
Db 2221 TATTTCCGCTTGAATACCTCTAGCTCTCTAGGTTATGTTGGTGTAGTATGATAGACAACAGG 2280
Qy 2281 GGATCTGTCAACCGAGGCTTAAATTTGAAAGGCGCTTTAAATATAAATATGAGTCAATA 2340
Db 2281 GGATCTGTCAACCGAGGCTTAAATTTGAAAGGCGCTTTAAATATAAATATGAGTCAATA 2340
Qy 2341 GAAATGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCAATGAC 2400
Db 2341 GAAATGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCAATGAC 2400
Qy 2401 TTAGATCGTGTGGGCATCCACCGCTGGTCTATGAGGAGTACCTCTCCCTGATGGCATTA 2460
Db 2401 TTAGATCGTGTGGGCATCCACCGCTGGTCTATGAGGAGTACCTCTCCCTGATGGCATTA 2460
Qy 2461 ATGACAGAGTCAGATATCTTCAAGGTTGCTATGCTGGGGCCCAAGTCACTCTGTGGATC 2520
Db 2461 ATGACAGAGTCAGATATCTTCAAGGTTGCTATGCTGGGGCCCAAGTCACTCTGTGGATC 2520
Qy 2521 TTTCTATGATAAGGATACACGGAACGTTATATGAGTCACTCTGACCAAGAAATGAACAGGC 2580
Db 2521 TTTCTATGATAAGGATACACGGAACGTTATATGAGTCACTCTGACCAAGAAATGAACAGGC 2580
Qy 2581 TATTTATAGGATCTGTGGGCATGCAAGCAAGAAAGTTCCCTCTGACCAAAATCGTTTA 2640
Db 2581 TATTTATAGGATCTGTGGGCATGCAAGCAAGAAAGTTCCCTCTGACCAAAATCGTTTA 2640
Qy 2641 CTGCTCTTATCATGTTTCTCTGGATGAGATGTCATTTTGCACATATCCAGTATATTAATCTG 2700
Db 2641 CTGCTCTTATCATGTTTCTCTGGATGAGATGTCATTTTGCACATATCCAGTATATTAATCTG 2700
Qy 2701 AGTTTATTTAGTGGGCTGGAAAGCATAATGATTTTACAGATCTATCTCTCAGGAGAGACAC 2760
Db 2701 AGTTTATTTAGTGGGCTGGAAAGCATAATGATTTTACAGATCTATCTCTCAGGAGAGACAC 2760
Qy 2761 AGCATAGAGTTCCTGAAATCGGGAGAACATTAATGAACTGCACTTTTGGCACTACTCTTCAA 2820
Db 2761 AGCATAGAGTTCCTGAAATCGGGAGAACATTAATGAACTGCACTTTTGGCACTACTCTTCAA 2820
Qy 2821 GAAAACTTTGGATCAGTATTTCTGCTCTTAAAGTGAATATTAATTTTGAACCTGTGTAGAAC 2880
Db 2821 GAAAACTTTGGATCAGTATTTCTGCTCTTAAAGTGAATATTAATTTTGAACCTGTGTAGAAC 2880
Qy 2881 TCTCTGGTATACCTGGCTATTTTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGA 2940
Db 2881 TCTCTGGTATACCTGGCTATTTTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGA 2940
Qy 2941 ATTGATCATCAATTTTGTATCTGCTGCAATGTAACATCTACTCTCTGAAAAATATAATGTGGTG 3000
Db 2941 ATTGATCATCAATTTTGTATCTGCTGCAATGTAACATCTACTCTCTGAAAAATATAATGTGGTG 3000
Qy 3001 CCATGAGGGGCTTACGGTTTGTGGTGTAGTAAATCTTAATCTTAAACCCCAATGTCTCAAAA 3060
Db 3001 CCATGAGGGGCTTACGGTTTGTGGTGTAGTAAATCTTAATCTTAAACCCCAATGTCTCAAAA 3060

Db 3001 CCATGCAGGGCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCCAATGCTCAAAA 3060
Qy 3061 TCAAAATGATACATATTCCTGAGAGACCCAGCAAAATACCAATAGATTAATTAATAAAAAA 3120
Db 3061 TCAAAATGATACATATTCCTGAGAGACCCAGCAAAATACCAATAGATTAATTAATAAAAAA 3120

RESULT 2
AF221634 3127 bp mRNA linear PRI 05-NOV-2000
LOCUS Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, complete cds.
DEFINITION AF221634.1 GI:11095187
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBLISHED
REFERENCE
AUTHORS
TITLE
JOURNAL
1. (bases 1 to 3127)
Abbott,C.A., Yu,D.M., Woollatt,E., Sutherland,G.R., McCaughan,G.W.
and Gorrell,M.D.
Cloning, expression and chromosomal localization of a novel human
dipeptidyl peptidase (DPP) IV homolog, DPP8
Eur. J. Biochem. 267 (20), 6140-6150 (2000)
2. (bases 1 to 3127)
Abbott,C.A., Yu,D., McCaughan,G.W. and Gorrell,M.D.
Direct Submission
Submitted (06-JAN-2000) A.W. Morrow Gastroenterology and Liver
Centre, Centenary Institute of Cell Biology and Cancer Medicine,
Locked Bag No.6, Newtown, Sydney, NSW 2042, Australia
Location/Qualifiers

FEATURES
source
1..3127
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q22"
1..3127
/gene="DPP8"
/gene="DPP8"
214..2862
/note="serine protease; member of prolyl oligopeptidase S9b
family belonging to enzyme clan SC; alternatively spliced
forms of this mRNA are found in AF221635, AF221636 and
AF221637"
/codon_start=1
/product="dipeptidyl peptidase 8"
/protein_id="AAC29766.1"
/db_xref="GI:11095188"
/translation="MAAMETQLGVEIFETADCEINIESQDRPKLEFFVYVWSQ
LKKLADIRKHGMKAPHDPMFKRNDPDGPHSDRIYILAMSGENRENTLYSEI
PKTINRAVLLSWKLLDFQATLDYGMYSREELLRRKRIQTVGIASYDYHQSG
TFLFQAGSLYHVDGPGQFTQOPLPNLVETSCFNIRMDPKLCPADPDWIAFIHNS
DIMISNITREERRLTYVHNLNMEEDARSAGVATFVLQEEFDRIYSGYWCPKAETT
PSGKILRIYLEENDESEVEI IHVTSFMLETRRADSFRPKGTANPKVTFPKSEIMI
DAEGRILDIVDKSLIOPFEILFEVGYIARAGMTPEKRYAWSILLDRSLQRLIVLIS
PELPEDVDVMERQRLIESVDSVTPLIYIETTDIWINIHDIHFVFPQSHHEEIEF
IFASECTGFRHLKYKITSILKESKYRSKSGGLPAPDFKCPKEBIAITSGEWEVLGR
HGSNIQVDERLVLYFEKTKDPSLEHLYVSVYVNPGEVTRLDTRGYSHSCISQHCD
FFTSKNQNPCHVLSLYKSSPEDDPTCKTKEFWATILDSAGPLPYDTPPEIFSPES
TTGTLTGMLYKPHDLQPGKKYPTVLFIYGGPQVQVNNRFGVKYFRFLANTLASLGV
VVDVNNGSCHRLGPEGAPKYMKGQIEIDQVEGLQYLASRYDFILDRVGIHGSY
GGYLSMALMORSIDIFRAIAGAPVLTWTFYDTGYTERYMGHPQNPQNGYVLGSVAMQ
AEKFPSPNPRLLLHGLFDENYHFAHTSILLFLVRAGKPYDLQIYQERHSIRVPES
GEHYELHLLHYQENLGSRIALKVI"

ORIGIN

Query Match 100.0%; Score 3120; DB 5; Length 3127;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGTGCTAAAGCCTCCGAGGCCAAGGCCGCTGCTACTCGCCCGCTGCTTCTTAGTCCG 60
Db 1 AAGTGCTAAAGCCTCCGAGGCCAAGGCCGCTGCTACTCGCCCGCTGCTTCTTAGTCCG 60
Qy 61 CGTTCCGCCCTGGGTGTGTACCCGCCGCCGCCGCGAGGAAGCCACTGCAACACGAGACCG 120
Db 61 CGTTCCGCCCTGGGTGTGTACCCGCCGCCGCCGCGAGGAAGCCACTGCAACACGAGACCG 120
Qy 121 GAGTGAGCGCGCGAGCAGATGAAGCGCGCGAGGCCGCCGCTCCATAGCGCAGTTCGGGACGG 180
Db 121 GAGTGAGCGCGCGAGCAGATGAAGCGCGCGAGGCCGCCGCTCCATAGCGCAGTTCGGGACGG 180
Qy 181 TCCGGCGCGCGCGCGGGGGAAGGAAATGCAACATGCGCAGCAGCAATGGGAAACAGAACAG 240
Db 181 TCCGGCGCGCGCGCGGGGGAAGGAAATGCAACATGCGCAGCAGCAATGGGAAACAGAACAG 240
Qy 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTCGAGGAGAAATATGGAATCACAGGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTCGAGGAGAAATATGGAATCACAGGATCGG 300
Qy 301 CCTAAATTCGAGCCCTTTTATGTTGAGCGGTATTCCTGAGTCACTTAAAGAGTCGTT 360
Db 301 CCTAAATTCGAGCCCTTTTATGTTGAGCGGTATTCCTGAGTCACTTAAAGAGTCGTT 360
Qy 361 GCCGATACCAGAAATATATCATGGCTACATGATGGCTAAGGCACCAATGGAATTCATGTTT 420
Db 361 GCCGATACCAGAAATATATCATGGCTACATGATGGCTAAGGCACCAATGGAATTCATGTTT 420
Qy 421 GTGAAGAGGATGATCCAGATGCGACCTCATTTCAGACAGAACTTATACCTTGGCCATGTCT 480
Db 421 GTGAAGAGGATGATCCAGATGCGACCTCATTTCAGACAGAACTTATACCTTGGCCATGTCT 480
Qy 481 GGTGAGAACAGAAATATACACTGTTTATTCGAAATCCCAAACTATCAATAGACGA 540
Db 481 GGTGAGAACAGAAATATACACTGTTTATTCGAAATCCCAAACTATCAATAGACGA 540
Qy 541 GCAGTCTTAATGCTCTCTTGGAGCCCTCTTTTGGATCTTTTTCAGGCAACACTTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTGGAGCCCTCTTTTGGATCTTTTTCAGGCAACACTTGGACTAT 600
Qy 601 GGAATGTAATTCGAGAAAGAAACTATTAAAGAGAAAGAAACGCATTCGAAACAGTCGGA 660
Db 601 GGAATGTAATTCGAGAAAGAAACTATTAAAGAGAAAGAAACGCATTCGAAACAGTCGGA 660
Qy 661 ATGTGCTTCTTACGATTATTCACCAAGGAAGTGGAAACATTTCTGTTTCAAGCCGCTAGTGA 720
Db 661 ATGTGCTTCTTACGATTATTCACCAAGGAAGTGGAAACATTTCTGTTTCAAGCCGCTAGTGA 720
Qy 721 ATTTATCACGTAAAGAGTGGAGGCCCAAGAGATTTTACGCAACAACTTTAAAGGCCCAAT 780
Db 721 ATTTATCACGTAAAGAGTGGAGGCCCAAGAGATTTTACGCAACAACTTTAAAGGCCCAAT 780
Qy 781 CTAGTGGAAACTAGTGTGCTCCCAACATACGGATGGATCCAAATTTATGCCCCGCTGATCCA 840
Db 781 CTAGTGGAAACTAGTGTGCTCCCAACATACGGATGGATCCAAATTTATGCCCCGCTGATCCA 840
Qy 841 GACTCGATTGCTTTTATATACATAGCAAGATTTTGGATATCTAAACATCGTAACCGAGAA 900
Db 841 GACTCGATTGCTTTTATATACATAGCAAGATTTTGGATATCTAAACATCGTAACCGAGAA 900
Qy 901 GAAAGGAGACTCACTTATGTGCAATAGCTAGCCAAACATCGAAGAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTATGTGCAACATAGCTAGCCAAACATCGAAGAGAGATGCCAGATCA 960
Qy 961 GCTGAGTTCGCTACCTTTGTTCTCCAGAGAGATTTGATAGATATTCGCTTATTCGCTGG 1020
Db 961 GCTGAGTTCGCTACCTTTGTTCTCCAGAGAGATTTGATAGATATTCGCTTATTCGCTGG 1020
Qy 1021 TGTCCAAAGCTGAACCACTCCCGAGTGGTGGTAAATTCCTTAGAATTCCTATGAGAA 1080
Db 1021 TGTCCAAAGCTGAACCACTCCCGAGTGGTGGTAAATTCCTTAGAATTCCTATGAGAA 1080
Qy 1081 AATGATGAATTCGAGGTGGAAATATTCATGTTACATCCCTATGTTGGAAACAGAGAGG 1140

[illegible]

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.									
1 Golz, S. Patent: WO 2005106021-A 1 10-NOV-2005; Bayer HealthCare AG (DE) Location/Qualifiers 1, -3127 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"									
source									
ORIGIN									
Query Match									
Best Local Similarity 98.4%; Score 3069; DB 2; Length 3127;									
Matches 3119; Conservative 99.9%; Pred. No. 0;									
Mismatches 1; Indels 0; Gaps 0;									
Qy	1	AGTGCTAAAGCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCTGCTTCTTAGTGCGG	60						
Db	1	AGTGCTAAAGCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCTGCTTCTTAGTGCGG	60						
Qy	61	CGTTGCGCGCTGGGTTGTCAACCGCGCCCGCCGAGGAAGCCACTGCAACCCAGGACCG	120						
Db	61	CGTTGCGCGCTGGGTTGTCAACCGCGCCCGCCGAGGAAGCCACTGCAACCCAGGACCG	120						
Qy	121	GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCCCGCTCCATAGCGCACGTCGCGACGG	180						
Db	121	GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCCCGCTCCATAGCGCACGTCGCGACGG	180						
Qy	181	TCCGGCGGGCGCGGGGGGAAGGAAATGCAACATGGCAGCAGCAATGGAACAGAACAG	240						
Db	181	TCCGGCGGGCGCGGGGGGAAGGAAATGCAACATGGCAGCAGCAATGGAACAGAACAG	240						
Qy	241	CTGGGTGTTCAGATATTTGAAACTGCGGACTGTGAGGAGATATTTGAATCAAGATCGG	300						
Db	241	CTGGGTGTTCAGATATTTGAAACTGCGGACTGTGAGGAGATATTTGAATCAAGATCGG	300						
Qy	301	CCTAAATTTGAGCCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGAGCTGCTT	360						
Db	301	CCTAAATTTGAGCCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGAGCTGCTT	360						
Qy	361	GCGGATACAGAAATATCATGGCTACATGATGGCTAAGGACCAACATGATTTCAATGTTT	420						
Db	361	GCGGATACAGAAATATCATGGCTACATGATGGCTAAGGACCAACATGATTTCAATGTTT	420						
Qy	421	GTGAGGAGATGATCCAGATGGACCTCATTCAGACAGAACTATTACCTTGGCATGTCT	480						
Db	421	GTGAGGAGATGATCCAGATGGACCTCATTCAGACAGAACTATTACCTTGGCATGTCT	480						
Qy	481	GGTGAGAACAGAGAAATAACATGTTTATTTCTGAAATTTCCCAAACTATCAATAGACA	540						
Db	481	GGTGAGAACAGAGAAATAACATGTTTATTTCTGAAATTTCCCAAACTATCAATAGACA	540						
Qy	541	GCAGTCTTAATGCTCTCTTGGAGCCTCTTTTGGATCTTTTCAAGCAACACTGGACTAT	600						
Db	541	GCAGTCTTAATGCTCTCTTGGAGCCTCTTTTGGATCTTTTCAAGCAACACTGGACTAT	600						
Qy	601	GGAAATGTAATCTCGAAGAGAAACTATTAAGAGAAAGAAACGCAATGGACAGTCGGA	660						
Db	601	GGAAATGTAATCTCGAAGAGAAACTATTAAGAGAAAGAAACGCAATGGACAGTCGGA	660						
Qy	661	ATTGCTCTTACGATATACCAAGGAGTGGACATTTCTGTTTCAAGCCGATGGA	720						
Db	661	ATTGCTCTTACGATATACCAAGGAGTGGACATTTCTGTTTCAAGCCGATGGA	720						
Qy	721	ATTATACAGTAAAGATGGAGGCGCACAGGATTTACGCAACACTTAAAGGCCAAT	780						
Db	721	ATTATACAGTAAAGATGGAGGCGCACAGGATTTACGCAACACTTAAAGGCCAAT	780						
Qy	781	CTAGTGGAAACTAGTGTGTTCCCAACATACAGGATGGATCCAAATTTATGCCCTCTCATCA	840						
Db	781	CTAGTGGAAACTAGTGTGTTCCCAACATACAGGATGGATCCAAATTTATGCCCTCTCATCA	840						

Qy	841	GACTGGATGCTTTTATATACATAGCAACGATATTTGGATATCTAACATCGTAACAGAGAA	900
Db	841	GACTGGATGCTTTTATATACATAGCAACGATATTTGGATATCTAACATCGTAACAGAGAA	900
Qy	901	GAAAGGAGACTCATTTATGTGCAACAATGAGCTAGCCAAACATGGAAGAGATCCAGATCA	960
Db	901	GAAAGGAGACTCATTTATGTGCAACAATGAGCTAGCCAAACATGGAAGAGATCCAGATCA	960
Qy	961	GCTGGAGTCGCTTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
Db	961	GCTGGAGTCGCTTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
Qy	1021	TGTCCTCAAAAGCTGAAACAACTCCCAAGTGGTGAATAATCTTTAGAAATCTTATATGAAGA	1080
Db	1021	TGTCCTCAAAAGCTGAAACAACTCCCAAGTGGTGAATAATCTTTAGAAATCTTATATGAAGA	1080
Qy	1081	ANTGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAACAAGGAGG	1140
Db	1081	ANTGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAACAAGGAGG	1140
Qy	1141	GCAGATTCATTCCTGTTATCTTAAACAGGTPACAGCAAAATCTTAAAGTCACTTTTAAGATG	1200
Db	1141	GCAGATTCATTCCTGTTATCTTAAACAGGTPACAGCAAAATCTTAAAGTCACTTTTAAGATG	1200
Qy	1201	TCAGAAATAATGATGATGCTGAGGAGGATCATAGATGTCTATAGATAAGGAACTAAATT	1260
Db	1201	TCAGAAATAATGATGATGCTGAGGAGGATCATAGATGTCTATAGATAAGGAACTAAATT	1260
Qy	1261	CAACCTTTTGAGATTTCTATTTGAAAGGAGTTGAATATATTGCGCAGAGCTGGATGGACTCT	1320
Db	1261	CAACCTTTTGAGATTTCTATTTGAAAGGAGTTGAATATATTGCGCAGAGCTGGATGGACTCT	1320
Qy	1321	GAGGAAAAATATGCTTGGTCCATCTTACTAGATCGTCCAGACTCGCTACAGATAGTG	1380
Db	1321	GAGGAAAAATATGCTTGGTCCATCTTACTAGATCGTCCAGACTCGCTACAGATAGTG	1380
Qy	1381	TTGATCTCACTCAATTTATTTATCCAGTAGAAGATGATGTTTGGAAAGCGAGACTC	1440
Db	1381	TTGATCTCACTCAATTTATTTATCCAGTAGAAGATGATGTTTGGAAAGCGAGACTC	1440
Qy	1441	ATTGAGTCAGTGCCTGATTCGTGAGCCACTAAATTTATCTATGAAGAAACAACAGACATC	1500
Db	1441	ATTGAGTCAGTGCCTGATTCGTGAGCCACTAAATTTATCTATGAAGAAACAACAGACATC	1500
Qy	1501	TGGATAAATATCCATGACATCTTCAATGTTTTTCCCAAAAGTCAGAGAGAGAAATTTGAG	1560
Db	1501	TGGATAAATATCCATGACATCTTCAATGTTTTTCCCAAAAGTCAGAGAGAGAAATTTGAG	1560
Qy	1561	TTTATTTTGGCTCTGAAATGCAAAAACAGGTTTCGTCAATTTATACAAATTTACATCTATT	1620
Db	1561	TTTATTTTGGCTCTGAAATGCAAAAACAGGTTTCGTCAATTTATACAAATTTACATCTATT	1620
Qy	1621	TTAAAGAAAGCAAAATATAAAGCATCCAGTGGTGGGCTGCTGCTCCAAAGTATTTCAAG	1680
Db	1621	TTAAAGAAAGCAAAATATAAAGCATCCAGTGGTGGGCTGCTGCTCCAAAGTATTTCAAG	1680
Qy	1681	TGTCCTATCAAAAGAGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTTGGCCGCGAT	1740
Db	1681	TGTCCTATCAAAAGAGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTTGGCCGCGAT	1740
Qy	1741	GGATCTAATATCAAGATTGATGAAGTCAAGAGGCTGGTATATTTTGAAGACACCAAGAC	1800
Db	1741	GGATCTAATATCAAGATTGATGAAGTCAAGAGGCTGGTATATTTTGAAGACACCAAGAC	1800
Qy	1801	TCCCTTTAGAGCATCACCTGTAGTACGTTACGTTAAATCTCTGAGAGGTTGACNAGG	1860
Db	1801	TCCCTTTAGAGCATCACCTGTAGTACGTTACGTTAAATCTCTGAGAGGTTGACNAGG	1860
Qy	1861	CTGACTGACCGTGGCTACTCACAATTTCTGCTGCACTAGTCAAGTGTGACTTCTTTATA	1920
Db	1861	CTGACTGACCGTGGCTACTCACAATTTCTGCTGCACTAGTCAAGTGTGACTTCTTTATA	1920
Qy	1921	AGTAAGTATAGTAAACAGAGAAATCCACATCTGTGTGCTCCCTTTTCAAGCTATCAAGTCTCT	1980

Db	1921	AGTAAGTATAGTAACACGAGAAATCCACACTGTGTGCTTTTACAAAGCTATCAAGTCT	1980
Qy	1981	GAAGATGACCCAACTTGGAAACAAAGAAATTTTGGGCGACCAATTTTGGATTCAGCAGGT	2040
Db	1981	GAAGATGACCCAACTTGGAAACAAAGAAATTTTGGGCGACCAATTTTGGATTCAGCAGGT	2040
Qy	2041	CCTCTCCCTGACTATCTACTCTCCAGAAATTTTCTTTTGAAAGTACTACTGGATTACA	2100
Db	2041	CCTCTCCCTGACTATCTACTCTCCAGAAATTTTCTTTTGAAAGTACTACTGGATTACA	2100
Qy	2101	TTGTATGGGATGCTCTACAGGCTCATCATCTACAGCTGGAAGAAATTCCTACTGTG	2160
Db	2101	TTGTATGGGATGCTCTACAGGCTCATCATCTACAGCTGGAAGAAATTCCTACTGTG	2160
Qy	2161	CTGTTCAATATATGGTGGTCTCAGGTGAGTGGTGAATATCGGTTTAAAGAGTCAAG	2220
Db	2161	CTGTTCAATATATGGTGGTCTCAGGTGAGTGGTGAATATCGGTTTAAAGAGTCAAG	2220
Qy	2221	TATTTCCGCTTGAATACCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGG	2280
Db	2221	TATTTCCGCTTGAATACCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGG	2280
Qy	2281	GGATCTCTGTCAACGAGGCTTAAATTTGAAGCGCTTAAATATATAAATGGGTCAATA	2340
Db	2281	GGATCTCTGTCAACGAGGCTTAAATTTGAAGCGCTTAAATATATAAATGGGTCAATA	2340
Qy	2341	GAATTTGACATCAGGTGGAAGACTCCATATCTAGCTTCTCGATATGATTTCAATGAC	2400
Db	2341	GAATTTGACATCAGGTGGAAGACTCCATATCTAGCTTCTCGATATGATTTCAATGAC	2400
Qy	2401	TTAGATCTGTGGGATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCATTA	2460
Db	2401	TTAGATCTGTGGGATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCATTA	2460
Qy	2461	ATGACAGGCTCAGATATCTTTCAAGGTTGCTATTTGCGGGCCCGACCTCTGTGATC	2520
Db	2461	ATGACAGGCTCAGATATCTTTCAAGGTTGCTATTTGCGGGCCCGACCTCTGTGATC	2520
Qy	2521	TTCTATGATACAGATACAGGACGTTATATGGGTCACCTGACAGATGAACAGGGC	2580
Db	2521	TTCTATGATACAGATACAGGACGTTATATGGGTCACCTGACAGATGAACAGGGC	2580
Qy	2581	TATTTACTTAGATCTGTGGCCATGCAAGCAGAAAGTTCCTCTGAACCAAATCGTTTA	2640
Db	2581	TATTTACTTAGATCTGTGGCCATGCAAGCAGAAAGTTCCTCTGAACCAAATCGTTTA	2640
Qy	2641	CTGCTCTTTACATGGTTTCTGGATGGAATGTCCATTTTGCACATACAGTATATCTG	2700
Db	2641	CTGCTCTTTACATGGTTTCTGGATGGAATGTCCATTTTGCACATACAGTATATCTG	2700
Qy	2701	AGTTTTTTTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACAC	2760
Db	2701	AGTTTTTTTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACAC	2760
Qy	2761	AGCATAAAGTTCCTGAAATCGGAGAACATTTATGAATGCAATCTTTTGCACACTACCTTCAA	2820
Db	2761	AGCATAAAGTTCCTGAAATCGGAGAACATTTATGAATGCAATCTTTTGCACACTACCTTCAA	2820
Qy	2821	GAAACCTTTGGATACGATTTGCTGCTCTAAAGTGATATAATTTTGAACCTGTGTAGAAC	2880
Db	2821	GAAACCTTTGGATACGATTTGCTGCTCTAAAGTGATATAATTTTGAACCTGTGTAGAAC	2880
Qy	2881	TCTCTGGTATACACTGGCTTATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGA	2940
Db	2881	TCTCTGGTATACACTGGCTTATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGA	2940
Qy	2941	ATTGATCATCACAATTTTGTATACCTGCCATGTAAACATCTACTCTCTGAAAAATTAATGGTG	3000
Db	2941	ATTGATCATCACAATTTTGTATACCTGCCATGTAAACATCTACTCTCTGAAAAATTAATGGTG	3000
Qy	3001	CCATGAGGGGTCTACGGTTTGTGTAGTAAATCTTAATACCTTAACCCACATCTCAAAA	3060
Db	3001	CCATGAGGGGTCTACGGTTTGTGTAGTAAATCTTAATACCTTAACCCACATCTCAAAA	3060
Qy	3061	TCAATGATACATATTTCTTGAGAGCCAGCAATACCATTAGAAATTTACTAAAAA	3120
Db	3061	TCAATGATACATATTTCTTGAGAGCCAGCAATACCATTAGAAATTTACTAAAAA	3120
RESULT 4			
AX342633			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
JOURNAL			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 3103; Conservative			
Qy	13	CTCCGAGGCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGCGCTTCGCGCT	72
Db	2	CTCCGAGGCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGCGCTTCGCGCT	61
Qy	73	GGTTGTCTCAACGCGCGCGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	132
Db	62	GGTTGTCTCAACGCGCGCGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	121
Qy	133	CGCAGATGAAGCGCGCGCGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	192
Db	122	CGCAGATGAAGCGCGCGCGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	181
Qy	193	CGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	252
Db	182	CGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	241
Qy	253	ATATTTGAACTGCGGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	312
Db	242	ATATTTGAACTGCGGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	301
Qy	313	CTTTTATGTTGAGCGGTTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	372
Db	302	CTTTTATGTTGAGCGGTTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	361
Qy	373	AAATATCATGGCTACATGATGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	432
Db	362	AAATATCATGGCTACATGATGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	421
Qy	433	GATCCAGATGGAGCTTATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	492
Db	422	GATCCAGATGGAGCTTATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	481
Qy	493	GAAATATACACTGTTTATTTCTGAAATTTCCAAATTTCCAAATTTCCAAATTTCCAA	552
Db	482	GAAATATACACTGTTTATTTCTGAAATTTCCAAATTTCCAAATTTCCAAATTTCCAA	541

Qy	553	CTCTCTTGGAAGCCCTCTTTTGGATCTTTTTCAGCGCAACACTGGACTATGGAAATGTTATCT	61
Db	542	CTCTCTTGGAAGCCCTCTTTTGGATCTTTTTCAGCGCAACACTGGACTATGGAAATGTTATCT	601
Qy	613	CGAGAAGAGAACTATTAAAGAGAAGAAAACGATTTGGAAACAGTCGGAAATGCTCTCTAC	672
Db	602	CGAGAAGAGAACTATTAAAGAGAAGAAAACGATTTGGAAACAGTCGGAAATGCTCTCTAC	661
Qy	673	GATTATCAACGAAGAGTGGAAACATTTCTGTGTTTCAAGCGGTAGTGGAAATTTATCAAGTA	732
Db	662	GATTATCAACGAAGAGTGGAAACATTTCTGTGTTTCAAGCGGTAGTGGAAATTTATCAAGTA	721
Qy	733	AAAGATGAGGAGCCACAAGATTTTACGCAACAACTTTAAGCCCAATCTAGTGGAAATCT	792
Db	722	AAAGATGAGGAGCCACAAGATTTTACGCAACAACTTTTACGCCCAATCTAGTGGAAATCT	781
Qy	793	AGTTGTGCCCAACATACGGATGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCT	852
Db	782	AGTTGTGCCCAACATACGGATGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCT	841
Qy	853	TTTATACATAGCAACGATTTTGGATCTCTAAACATCGTAAACAGAGAAGAAAGGAGACTC	912
Db	842	TTTATACATAGCAACGATTTTGGATCTCTAAACATCGTAAACAGAGAAGAAAGGAGACTC	901
Qy	913	ACTTATGTCACATAGCTAGCTAGCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGT	972
Db	902	ACTTATGTCACATAGCTAGCTAGCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGT	961
Qy	973	ACCTTTGTTCTCCMAGAGAAATTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAAGCT	1032
Db	962	ACCTTTGTTCTCCMAGAGAAATTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAAGCT	1021
Qy	1033	GAACAACATCCCAGTGGTGGTAAAAATCTTAGAAATCTTATGAAGAAAATCATCAATCT	1092
Db	1022	GAACAACATCCCAGTGGTGGTAAAAATCTTAGAAATCTTATGAAGAAAATCATCAATCT	1081
Qy	1093	GAGGTGAAAATTTATCATGTTACATCCCTCATGTTGGAAACAAGGAGGCGCAGATTCATTC	1152
Db	1082	GAGGTGAAAATTTATCATGTTACATCCCTCATGTTGGAAACAAGGAGGCGCAGATTCATTC	1141
Qy	1153	CGTTATCCTAAAAACAGGTACAGCAAACTCTTAAAGTCACTTTTAAAGATGTCAGAAATATG	1212
Db	1142	CGTTATCCTAAAAACAGGTACAGCAAACTCTTAAAGTCACTTTTAAAGATGTCAGAAATATG	1201
Qy	1213	ATTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAACATAATTCACCTTTTGAG	1272
Db	1202	ATTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAACATAATTCACCTTTTGAG	1261
Qy	1273	ATTCTATTTGAAAGAGTTTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAT	1332
Db	1262	ATTCTATTTGAAAGAGTTTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAT	1321
Qy	1333	GCTTGGTCCATCCTACTAGATGCTCCCGAGCTCGCTACAGATAGTGTGATCTCACT	1392
Db	1322	GCTTGGTCCATCCTACTAGATGCTCCCGAGCTCGCTACAGATAGTGTGATCTCACT	1381
Qy	1393	GAATATTTATCCCAAGTAAAGATGATTTATGGAANGCAGAGACTCATTTAGTGTCAAGT	1452
Db	1382	GAATATTTATCCCAAGTAAAGATGATTTATGGAANGCAGAGACTCATTTAGTGTCAAGT	1441
Qy	1453	CCTGATTTCTGAGCGCCATTAATTTATCTATGAAGAAAACACAGACATCTGGATAAATATC	1512
Db	1442	CCTGATTTCTGAGCGCCATTAATTTATCTATGAAGAAAACACAGACATCTGGATAAATATC	1501
Qy	1513	CATGACATCTTTCATGTTTTTCCCACAAAGTCAAGAGAGAAATTTAGTGTATTTTGGC	1572
Db	1502	CATGACATCTTTCATGTTTTTCCCACAAAGTCAAGAGAGAAATTTAGTGTATTTTGGC	1561
Qy	1573	TCTGAAATGCAAAACAGGTTCCCGTCATTTATACAAAATTTACATCTATTATTAAGGAAAGC	1632
Db	1562	TCTGAAATGCAAAACAGGTTCCCGTCATTTATACAAAATTTACATCTATTATTAAGGAAAGC	1621

QY	1633	AAATATAAACGATCCAGTGGTGGCTGCTGCTCCAAGTGATTTCAAAGTGTCTTATCAAA	1692
DB	1622	AAATATAAACGATCCAGTGGTGGCTGCTGCTCCAAGTGATTTCAAAGTGTCTTATCAAA	1681
QY	1693	GAGGAGATAGCAATTAACGATGGTGAAATGGGAAGTTCTTGCCGGCATGATCTAATATC	1752
DB	1682	GAGGAGATAGCAATTAACGATGGTGAAATGGGAAGTTCTTGCCGGCATGATCTAATATC	1741
QY	1753	CAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTCAAGGCAACCAAGACTCCCTTTTAGAG	1812
DB	1742	CAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTCAAGGCAACCAAGACTCCCTTTTAGAG	1801
QY	1813	CATCACTCTGACGTAGTCAGTTACGTAAATCCTCGAGAGTGACAGAGCTGACGACCGT	1872
DB	1802	CATCACTCTGACGTAGTCAGTTACGTAAATCCTCGAGAGTGACAGAGCTGACGACCGT	1861
QY	1873	GGCTACTCACATTTCTTGCTGCATCAGTCAGCACGTGTGACTCTTTATTAAGTAAGTATAGT	1932
DB	1862	GGCTACTCACATTTCTTGCTGCATCAGTCAGCACGTGTGACTCTTTATTAAGTAAGTATAGT	1921
QY	1933	AACCCAGAGATCCACACTGTGTGTCCTTTACAGCTATCAAGTCTGTAAGATGACCCA	1992
DB	1922	AACCCAGAGATCCACACTGTGTGTCCTTTACAGCTATCAAGTCTGTAAGATGACCCA	1981
QY	1993	ACTTGCACAAACAAAGGAATTTTGGCCACCATTTTGGATTCACAGAGTCTCTTCTCGAC	2052
DB	1982	ACTTGCACAAACAAAGGAATTTTGGCCACCATTTTGGATTCACAGAGTCTCTTCTCGAC	2041
QY	2053	TATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATG	2112
DB	2042	TATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATG	2101
QY	2113	CTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTGCTTTGATATAT	2172
DB	2102	CTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTGCTTTGATATAT	2161
QY	2173	GGTGGTCTCAGGTGCGAGTTGGTGAATAATTCGGTTTAAAGGAGTCAAAGTATTTCCGCTTG	2232
DB	2162	GGTGGTCTCAGGTGCGAGTTGGTGAATAATTCGGTTTAAAGGAGTCAAAGTATTTCCGCTTG	2221
QY	2233	AATACCTTAGCCTCTTAGTGTATGTGGTTGTAGTGATAGACAACAGGGGATCTGTCAC	2292
DB	2222	AATACCTTAGCCTCTTAGTGTATGTGGTTGTAGTGATAGACAACAGGGGATCTGTCAC	2281
QY	2293	CGAGGGCTTAAATTTGAAGCGGCTTTAAATATATAATATAAATGGGTCAAATAGAAATGACGAT	2352
DB	2282	CGAGGGCTTAAATTTGAAGCGGCTTTAAATATATAATATAAATGGGTCAAATAGAAATGACGAT	2341
QY	2353	CAGGTGAAAGGACTCCAATATCTTAGTCTTCGATATGATTTTCATTGACTTAGATCGGTG	2412
DB	2342	CAGGTGAAAGGACTCCAATATCTAGTCTTCGATATGATTTTCATTGACTTAGATCGGTG	2401
QY	2413	GGCATCCACGGCTGGTCTTATCGAGGATACCTCTCCCTGATGGCATTAATCGAGAGTCA	2472
DB	2402	GGCATCCACGGCTGGTCTTATCGAGGATACCTCTCCCTGATGGCATTAATCGAGAGTCA	2461
QY	2473	GATATCTTCAGGGTTCCTATTCGTGGGCCCGAGTCACTCTGTGGATCTTCTATGATACA	2532
DB	2462	GATATCTTCAGGGTTCCTATTCGTGGGCCCGAGTCACTCTGTGGATCTTCTATGATACA	2521
QY	2533	GGATACACGGAACTTATATGGGTCACTCGATGACAGAAATGAACAGGGCTATTACTTAGGA	2592
DB	2522	GGATACACGGAACTTATATGGGTCACTCGATGACAGAAATGAACAGGGCTATTACTTAGGA	2581
QY	2593	TCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACAT	2652
DB	2582	TCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACAT	2641
QY	2653	GGTTTCCTCGATGAGAAATGTCATTTTGACATACCAAGTATATTACTAGGTTTTTTTAGTG	2712
DB	2642	GGTTTCCTCGATGAGAAATGTCATTTTGACATACCAAGTATATTACTAGGTTTTTTTAGTG	2701
QY	2713	AGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATTAAGGTT	2772

Db	2702	AGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAGAGTT	2761	
Qy	2773	CCTGAATCGGAGAACATTTATGAATCTCTTTTGGCACTACTCTTCAAGAAAACTTTGGA	2832	
Db	2762	CCTGAATCGGAGAACATTTATGAATCTCTTTTGGCACTACTCTTCAAGAAAACTTTGGA	2821	
Qy	2833	TCAGTATTGCTGCTCTTAAAGTGATATAATTTGACCTGTGTAGAACTCTCTGTATAC	2892	
Db	2822	TCAGTATTGCTGCTCTTAAAGTGATATAATTTGACCTGTGTAGAACTCTCTGTATAC	2881	
Qy	2893	ACTGGCTATTTAACCAATGAGGAGGTTTAAATCAACAGAAAAACAGAAATGATCATCAC	2952	
Db	2882	ACTGGCTATTTAACCAATGAGGAGGTTTAAATCAACAGAAAAACAGAAATGATCATCAC	2941	
Qy	2953	ATTTTGATACCTGCGCATGTAACTACTCTCTGAAATAATTTGCTGCGCATGCGGGGT	3012	
Db	2942	ATTTTGATACCTGCGCATGTAACTACTCTCTGAAATAATTTGCTGCGCATGCGGGGT	3001	
Qy	3013	CTACGGTTTGGTGTAGTAATCTAATACCTTTAACCCACACATGCTCAAAATCAAATGATACA	3072	
Db	3002	CTACGGTTTGGTGTAGTAATCTAATACCTTTAACCCACACATGCTCAAAATCAAATGATACA	3061	
Qy	3073	TATTCCTGAGAGCCAGCAATACCAATAGAAATTTACTAAAAAAA	3117	
Db	3062	TATTCCTGAGAGCCAGCAATACCAATAGAAATTTACTAAAAAAA	3106	
RESULT 5				
AX354793		3143 bp	DNA	linear
LOCUS	AX354793			PAT 06-FEB-2002
DEFINITION	Sequence 1 from Patent WO0179473.			
ACCESSION	AX354793			
VERSION	AX354793.1	GI:18619526		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
	Hominidae; Homo.			
REFERENCE	1			
AUTHORS	Meyers, R.A. and Williamson, M.			
TITLE	21953, a human prollyl oligopeptidase family member and uses thereof			
JOURNAL	Patent: WO 0179473-A 1 25-OCT-2001:			
	Millennium Pharmaceuticals, Inc. (US)			
FEATURES	Location/Qualifiers			
source	1..3143			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
	229..2877			
	/note="unnamed protein product"			
	/codon_start=1			
	/protein_id="CAD2290.1"			
	/db_xref="GI:18619527"			
CDS				
	/translation="MAAMETEQLGVRIPTADCEENIESQDRPKLEPPYVERYSWSQ			
	LKLLADTRKHGYMAKAPHDPMFVKENDPDGPHSDRIYLYLWAMGENENTLFSYSEI			
	PKTINRAAVMLSWKPLDLQATLDYGMTSREELLREKRIGTGVGLASDYHQSG			
	TEFLQAGSIVYHDKGGPQGTQQLPRLNIVTSVCPNIRMDPKLCPADPDWAFIHSN			
	DIWIGNVTRERRLTYVHNELANMEEDARSAGVATFVLQEEEDFYSGYVWCPKAEIT			
	PSGKILRIIYENDESEVETIIVHTSPMLETRADSPRYKPTGTANPKVTPFWSIMI			
	DAEGRILIDVDEKELQPEELFEGVEVIARAGTPEGKYAWSILLDRSOTRLQIVLIS			
	PELFIPEDDWEROELIESVPDSVTPLLIYEETDIIWINHDI FHVFPQSHHEEIEF			
	IFASCKTGFRHLYKITSILKESIKKSSGLPAPDFKCPKEEIAITSGSEWVLGR			
	HGSNIQVDEVRRLVYFEGKDSPLEHLYVYVYVNPGEVTRLDYDRGYSHSCCSQHD			
	FTISKYSNKPCHVCVLYKLSPPEDPTCKTFWATILDSAGRPDPTTPPIFSFES			
	TTGTFLYGLYKHPDLOPKKPYVLFYGGPQVQLVNNRFKGVKFRYLRNLTASLYV			
	VVVIDNRSCKGLKFEKAFKVMGOLEIDDOVEGLQYLSRVDFDLDKRVGHGWSY			
	GGTLSMALMQRSDFPRVAIAGAPVTLWIFDYGYTERYMGHPDQEQGYLGSVAMQ			
	AKKFPSEPNRLLHLGLDENVHFAHTSILLISFLVRACKPYDLQIYQERHSIRVPS			
	GBHYELHLLHYLQENLIGSRIALKVI"			

ORIGIN

Query Match	94.8%;	Score 2957;	DB 2;	Length 3143;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 3107;	Conservative	0;	Mismatches	3;
			Indels	0;
			Gaps	0;
Qy	11	GCCTCCGAGGCCAAGGCGCTGCTACTGCGCGCGTCTTCTTAGTGCGCGCTTCGCGCG	70	
Db	26	GCCTCCGAGGCCAAGGCGCTGCTACTGCGCGCGTCTTCTTAGTGCGCGCTTCGCGCG	85	
Qy	71	CTGGTGTGTCAACCGCGCGCGCGCGCGAGGAGCACTGCAACAGGACCGAGAGTGAGGC	130	
Db	86	CTGGTGTGTCAACCGCGCGCGCGCGAGGAGCACTGCAACAGGAGCGAGTGAGGC	145	
Qy	131	GGCGCAGCATGAAGCGCGCGCGCGCGCGAGGCGCGCTCCATAGCGCAGCTCGGAGCGG	190	
Db	146	GGCGCAGCATGAAGCGCGCGCGCGCGAGGCGCGCTCCATAGCGCAGCTCGGAGCGG	205	
Qy	191	GGCGGGGGGAGGAAATGCAACATGCGCAGCAATGGAACAGAAACAGAACTGGGTGTG	250	
Db	206	GGCGGGGGGAGGAAATGCAACATGCGCAGCAATGGAACAGAAACAGAACTGGGTGTG	265	
Qy	251	AGATATTTTGAACCTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGGCTTAATTGG	310	
Db	266	AGATATTTTGAACCTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGGCTTAATTGG	325	
Qy	311	AGCCTTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAAAGCTTTCGCGATACCA	370	
Db	326	AGCCTTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAAAGCTTTCGCGATACCA	385	
Qy	371	GAAATATCATGGCTACATGATGCTAAGGCACACATGATTTTCATGTTTGTGAAGAGA	430	
Db	386	GAAATATCATGGCTACATGATGCTAAGGCACACATGATTTTCATGTTTGTGAAGAGA	445	
Qy	431	ATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCATGTCTGGTGAGAA	490	
Db	446	ATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCATGTCTGGTGAGAA	505	
Qy	491	GAGAAAAACACTGTTTATTTCTGAAATTTCCAAAACATCAATAGACGAGCTCTTAA	550	
Db	506	GAGAAAAACACTGTTTATTTCTGAAATTTCCAAAACATCAATAGACGAGCTCTTAA	565	
Qy	551	TGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGGACTATGGAATGTATT	610	
Db	566	TGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGGACTATGGAATGTATT	625	
Qy	611	CTCGAGAGAGAACTATTAAAGAGAGAAAAACGATTTGGAAACAGTTCGGAATTCCTT	670	
Db	626	CTCGAGAGAGAACTATTAAAGAGAGAAAAACGATTTGGAAACAGTTCGGAATTCCTT	685	
Qy	671	ACGATTTATCAACAGGAGTGGAACTTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGC	730	
Db	686	ACGATTTATCAACAGGAGTGGAACTTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGC	745	
Qy	731	TAAAGATCGAGGCGCAACAGGATTTAGCAACAACTTTTAAAGCGCCCAATCTAGTGGAA	790	
Db	746	TAAAGATCGAGGCGCAACAGGATTTAGCAACAACTTTTAAAGCGCCCAATCTAGTGGAA	805	
Qy	791	CTAGTTTCTCCCAACATACGATTTGATTTCCAAATTTATGCCCCGTGATCCAGACTGGATTG	850	
Db	806	CTAGTTTCTCCCAACATACGATTTGATTTCCAAATTTATGCCCCGTGATCCAGACTGGATTG	865	
Qy	851	CTTTTATACATAGCAACGATTTTGGATTTCTAATCTGTAACCAAGAGAGAGAGAGAC	910	
Db	866	CTTTTATACATAGCAACGATTTTGGATTTCTAATCTGTAACCAAGAGAGAGAGAGAC	925	
Qy	911	TCATTTATGCAATAGCTAGCAACATGGAAGAGATGCGAGATCAGCTGGATCG	970	
Db	926	TCATTTATGCAATAGCTAGCTAGCAACATGGAAGAGATGCGAGATCAGCTGGATCG	985	
Qy	971	CTACCTTTTGTCTTCAAGAGAAATTTGATAGATATTCCTGGCTATTTGGTGTCTCAAAG	1030	
Db	986	CTACCTTTTGTCTTCAAGAGAAATTTGATAGATATTCCTGGCTATTTGGTGTCTCAAAG	1045	
Qy	1031	CTGAAACAACTCCAGTGTGTGTAAATTTCTTAGAATTTCTATATGAAGAAAAATGATGAAT	1090	

Db 1046 |||||CTGAACAACCTCCAGTGGTGAATTTCTAGAAATCTATATGAAGAAATATGAAT 1105
Qy 1091 CTGAGTGGGAATATTTCATGTTACATCCCTATGTTGGAAAACAGAGAGGCGAGATTCAT 1150
Db 1106 CTGAGGTGGAAATATTTCATGTTACATCCCTATGTTGGAAACAGAGAGGCGAGATTCAT 1165
Qy 1151 TCCGTTATCTCTAAACAGGTACAGCAATCTTAAAGTCACCTTTTAAAGATGTCAGAAATAA 1210
Db 1166 TCCGTTATCTCTAAACAGGTACAGCAATCTTAAAGTCACCTTTTAAAGATGTCAGAAATAA 1225
Qy 1211 TGATGATGCTGAAAGAGGATCATAGATGTCATAGATAAGAACTAAATCAAACCTTTTG 1270
Db 1226 TGATGATGCTGAAAGAGGATCATAGATGTCATAGATAAGAACTAAATCAAACCTTTTG 1285
Qy 1271 AGATTCATTTGAAAGAGTGAATATATGTCAGAGCTGGATGACCTCTGAGGGAAAT 1330
Db 1286 AGATTCATTTGAAAGAGTGAATATATGTCAGAGCTGGATGACCTCTGAGGGAAAT 1345
Qy 1331 ATGCTTTGGTCCATCTCTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCAC 1390
Db 1346 ATGCTTTGGTCCATCTCTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCAC 1405
Qy 1391 CTGAATTTATTTCCAGTGAAGATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCAG 1450
Db 1406 CTGAATTTATTTCCAGTGAAGATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCAG 1465
Qy 1451 TGCTGATTTCTGTAGCGCCACTTAATTTATCTATGAAGAAACAACAGACATCTGGATATA 1510
Db 1466 TGCTGATTTCTGTAGCGCCACTTAATTTATCTATGAAGAAACAACAGACATCTGGATATA 1525
Qy 1511 TCCATGACATCTTTTCATGTTTCCCAAGTGCAGAGGAGAAATGAGTTATTTTG 1570
Db 1526 TCCATGACATCTTTTCATGTTTCCCAAGTGCAGAGGAGAAATGAGTTATTTTG 1585
Qy 1571 CCTCTGAATGGAACAGGTTCCGTCATTTATACAAAATATACATCTATTTTAAAGGAAA 1630
Db 1586 CCTCTGAATGGAACAGGTTCCGTCATTTATACAAAATATACATCTATTTTAAAGGAAA 1645
Qy 1631 GCAATATTAACCATCCAGTGGTGGCTGCTGCTCCAGTGATTTCAAGTGATCTCTATCA 1690
Db 1646 GCAATATTAACCATCCAGTGGTGGCTGCTGCTCCAGTGATTTCAAGTGATCTCTATCA 1705
Qy 1691 AAGAGGAGATGCAATACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATA 1750
Db 1706 AAGAGGAGATGCAATACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATA 1765
Qy 1751 TCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCAACCAAGACTCCCTTTAG 1810
Db 1766 TCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCAACCAAGACTCCCTTTAG 1825
Qy 1811 AGCATCACCTGTAGTACAGTTACGTAAATCTCGAGAGGTGACAAGGCTGACTGACC 1870
Db 1826 AGCATCACCTGTAGTACAGTTACGTAAATCTCGAGAGGTGACAAGGCTGACTGACC 1885
Qy 1871 GTGGTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTATA 1930
Db 1886 GTGGTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTATA 1945
Qy 1931 GTAACCAAGAAATCCACACTGTGTGCTCCCTTTACAGCTATCAAGTCTCTGAGATGACC 1990
Db 1946 GTAACCAAGAAATCCACACTGTGTGCTCCCTTTACAGCTATCAAGTCTCTGAGATGACC 2005
Qy 1991 CAACCTTGCAAAACAAAGAAATTTTGGGCCACCAATTTGGATTCAGCAGTCTCTCTCTG 2050
Db 2006 CAACCTTGCAAAACAAAGAAATTTTGGGCCACCAATTTGGATTCAGCAGTCTCTCTCTG 2065
Qy 2051 ACTATACCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTACATGTTAGTGGGA 2110
Db 2066 ACTATACCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTACATGTTAGTGGGA 2125
Qy 2111 TGCTCTCAAGCTCATGATCTACAGCTGGAAAGAAATATCTCTATGTGCTGTTCAAT 2170

Db 2126 TGCTCTCAAGAGCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTGCTGTTCATAT 2185
Qy 2171 ATGGTGGTCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTTCCGCT 2230
Db 2186 ATGGTGGTCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTTCCGCT 2245
Qy 2231 TGAATACCCCTAGCTCTCTAGGTTATGTTGGTTGTAGTGATAGACAAACAGGGGATCTCTGTC 2290
Db 2246 TGAATACCCCTAGCTCTCTAGGTTATGTTGGTTGTAGTGATAGACAAACAGGGGATCTCTGTC 2305
Qy 2291 ACCGAGGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATCGGTCAAAATAGAAATTCACG 2350
Db 2306 ACCGAGGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATCGGTCAAAATAGAAATTCACG 2365
Qy 2351 ATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTTGACTTAGATCGTG 2410
Db 2366 ATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTTGACTTAGATCGTG 2425
Qy 2411 TGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCGTGATGGCAATTAATGCGAGGT 2470
Db 2426 TGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCGTGATGGCAATTAATGCGAGGT 2485
Qy 2471 CAGATATCTTCAAGGTTGCTATTTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGATA 2530
Db 2486 CAGATATCTTCAAGGTTGCTATTTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGATA 2545
Qy 2531 CAGGATACAGGAAGCTTATGAGGTACCTCTGACCAAGATGAACAGGGCTATTTACTTAG 2590
Db 2546 CAGGATACAGGAAGCTTATGAGGTACCTCTGACCAAGATGAACAGGGCTATTTACTTAG 2605
Qy 2591 GATCTGTGGCCATGCAAGAGAAAGTTCCCTCTGAACCAAAATCGTTTACTCTCTTAC 2650
Db 2606 GATCTGTGGCCATGCAAGAGAAAGTTCCCTCTGAAACCAAAATCGTTTACTCTCTTAC 2665
Qy 2651 ATGGTTTCTGTGATGAGAAATGTCATTTTGACATACCACTATATTAAGTGTGTTTTAG 2710
Db 2666 ATGGTTTCTGTGATGAGAAATGTCATTTTGACATACCACTATATTAAGTGTGTTTTAG 2725
Qy 2711 TGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAGCATTAAGAG 2770
Db 2726 TGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAGCATTAAGAG 2785
Qy 2771 TTCTCTGAATCGGGGAAACATTTATGAACCTGCACTTTTTTGCACTACCTTCAAGAAAACTTG 2830
Db 2786 TTCTCTGAATCGGGGAAACATTTATGAACCTGCACTTTTTTGCACTACCTTCAAGAAAACTTG 2845
Qy 2831 GATCAGTATTTGCTCTTAAAGTGATATAATTTTGACCTGTGTAGAACCTCTCTGAT 2890
Db 2846 GATCAGTATTTGCTCTTAAAGTGATATAATTTTGACCTGTGTAGAACCTCTCTGAT 2905
Qy 2891 ACACCTGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACAGAAATGATCATC 2950
Db 2906 ACACCTGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACAGAAATGATCATC 2965
Qy 2951 ACATTTTGAATACCTGCAATGTAACATCTACTCTGAAATAAATGTTGGTCCCATGAGGG 3010
Db 2966 ACATTTTGAATACCTGCAATGTAACATCTACTCTGAAATAAATGTTGGTCCCATGAGGG 3025
Qy 3011 GTCTACGGTTTGTGTAGTAAATCTAATACCTTAACCCCATGCTCAAAATCAAAATGATA 3070
Db 3026 GTCTACGGTTTGTGTAGTAAATCTAATACCTTAACCCCATGCTCAAAATCAAAATGATA 3085
Qy 3071 CATATTCCTGAGAGACCCAGCAATACATTAAGAAATTTACTTAAAAAATAAAAA 3120
Db 3086 CATATTCCTGAGAGACCCAGCAATACATTAAGAAATTTACTTAAAAAATAAAAA 3135

RESULT 6
ARG31283
LOCUS
DEFINITION
ACCESSION
VERSION

ARG31283
Sequence 12 from patent US 6844180.
ARG31283
VERSION
4829 bp DNA linear PAT 14-FEB-2005
GI:59770927

KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 4829)
AUTHORS	Oi, S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.
TITLE	Serine protease genes related to DPP1V
JOURNAL	Patent: US 6841180-A 12 18-JAN-2005;
	Ferring BV,;
FEATURES	Location/Qualifiers
source	1..4829
	/organism="unknown"
	/mol_type="genomic DNA"
ORIGIN	
	Query Match 92.6%; Score 2889; DB 2; Length 4829;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
Qy	1 AAGTGTAAAGCCTCGAGGCCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGGCG 60
Db	1 AAGTGTAAAGCCTCGAGGCCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGGCG 60
Qy	61 CGTTCCGCGCTGGTGTGTCACCGCGCGCGCGCGAGGAGCCACTGGCAACGAGACCG 120
Db	61 CGTTCCGCGCTGGTGTGTCACCGCGCGCGCGCGAGGAGCCACTGGCAACGAGACCG 120
Qy	121 GAGTGAAGCGCGGAGCATGAAGCGCGCGAGCCGCTCCATAGCGCACGTGGGACGG 180
Db	121 GAGTGAAGCGCGGAGCATGAAGCGCGCGAGCCGCTCCATAGCGCACGTGGGACGG 180
Qy	181 TCCGGCGCGCGCGGAGGAAATGCAACATGCGCAGCAGCAATGGAAACAGAACAG 240
Db	181 TCCGGCGCGCGCGGAGGAAATGCAACATGCGCAGCAGCAATGGAAACAGAACAG 240
Qy	241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGATATTTGAATCACAGATCGG 300
Db	241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGATATTTGAATCACAGATCGG 300
Qy	301 CCTAAATGAGCCTTTTATGTTGAGCGGTATTTCTGGAGTCACTTTAAAGAGCTGCTT 360
Db	301 CCTAAATGAGCCTTTTATGTTGAGCGGTATTTCTGGAGTCACTTTAAAGAGCTGCTT 360
Qy	361 GCCGATACAGAAATATCATGGCTACATGATGGCTAAGGCCACCAATGATTTTCAATGTTT 420
Db	361 GCCGATACAGAAATATCATGGCTACATGATGGCTAAGGCCACCAATGATTTTCAATGTTT 420
Qy	421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTCTATTACCTTGGCATGTCT 480
Db	421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTCTATTACCTTGGCATGTCT 480
Qy	481 GGTGAACACAGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540
Db	481 GGTGAACACAGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540
Qy	541 GCAGTCTTAATGCTCTCTTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db	541 GCAGTCTTAATGCTCTCTTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Qy	601 GGAATGTATTCTGAGAGAGAACTATTAGAGAAAGAAACGCAATGGAAACAGTCGGA 660
Db	601 GGAATGTATTCTGAGAGAGAACTATTAGAGAAAGAAACGCAATGGAAACAGTCGGA 660
Qy	661 ATTGCTCTTACGATTATCACCAAGGAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGA 720
Db	661 ATTGCTCTTACGATTATCACCAAGGAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGA 720
Qy	721 ATTATACAGTAAAGATGGAGGCCACCAAGGATTTACGCAACACCTTTAAGGCCCAAT 780
Db	721 ATTATACAGTAAAGATGGAGGCCACCAAGGATTTTACGCAACACCTTTAAGGCCCAAT 780
Qy	781 CTAGTGGAAACTAGTTGTCCCAACATACGGATGGATCCAAATTTATGCCCGCTGATCCA 840

Db	781 CTAGTGGAAACTAGTTGTCCCAACATACGGATGGATCCAAATTTATGCCCTGCTGATCCA 840
Qy	841 GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAAATCGTAACACAGAGAA 900
Db	841 GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAAATCGTAACACAGAGAA 900
Qy	901 GAAAGGAGACTCAGTTATGTGACAAATGAGCTAGCCAAATGGAAGAAAGATGCGAGATCA 960
Db	901 GAAAGGAGACTCAGTTATGTGACAAATGAGCTAGCCAAATGGAAGAAAGATGCGAGATCA 960
Qy	961 GCTGGAGTCGCTA CTTTGTCTTCTCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Db	961 GCTGGAGTCGCTA CTTTGTCTTCTCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Qy	1021 TGTCCAAAGAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA 1080
Db	1021 TGTCCAAAGAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA 1080
Qy	1081 AATGATGAATCTGAGGTGGAATTTATCATGTTACATCCCTATGTTGGAAACAGAGAGG 1140
Db	1081 AATGATGAATCTGAGGTGGAATTTATCATGTTACATCCCTATGTTGGAAACAGAGAGG 1140
Qy	1141 GCAGATTCAATTCGTTATCCTTAAACAGGTACAGCAATCCTAAAGTCACCTTTTAAAGATG 1200
Db	1141 GCAGATTCAATTCGTTATCCTTAAACAGGTACAGCAATCCTAAAGTCACCTTTTAAAGATG 1200
Qy	1201 TCAGAAATATGATGATGCTGTAAGGAGGATCATAGATGTCTAGATTAAGAACTAATTT 1260
Db	1201 TCAGAAATATGATGATGCTGTAAGGAGGATCATAGATGTCTAGATTAAGAACTAATTT 1260
Qy	1261 CAACCTTTTCAGATTTCTATTTGAAAGGATTTGAATATATTTGCCAGAGCTGGATGGACTCCT 1320
Db	1261 CAACCTTTTCAGATTTCTATTTGAAAGGATTTGAATATATTTGCCAGAGCTGGATGGACTCCT 1320
Qy	1321 GAGGAAAAATATGCTTTGGTCCATCTCTACTAGATCGCTCCACAGCTCGCTTACAGATAGTG 1380
Db	1321 GAGGAAAAATATGCTTTGGTCCATCTCTACTAGATCGCTCCACAGCTCGCTTACAGATAGTG 1380
Qy	1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTC 1440
Db	1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTC 1440
Qy	1441 ATTGAGTCAGTCGCTGATTTCTGTGACGCCACTAATTTCTATGGAAGAAACACAGACATC 1500
Db	1441 ATTGAGTCAGTCGCTGATTTCTGTGACGCCACTAATTTCTATGGAAGAAACACAGACATC 1500
Qy	1501 TGGATAAATATCCATGACATCTTTTCAATGTTTTTCCCAAAAGTCACGAAGAGGAAATTTGAG 1560
Db	1501 TGGATAAATATCCATGACATCTTTTCAATGTTTTTCCCAAAAGTCACGAAGAGGAAATTTGAG 1560
Qy	1561 TTTATTTTGTGCTCTGAAATGCAAAACAGGTTTCCGTCATTTTATACAAAATTACATCTATT 1620
Db	1561 TTTATTTTGTGCTCTGAAATGCAAAACAGGTTTCCGTCATTTTATACAAAATTACATCTATT 1620
Qy	1621 TTTAAAGGAAGCAAAATATAACGATCGATGGGCTGCTGCTCCAAGTGAATTTCAAG 1680
Db	1621 TTTAAAGGAAGCAAAATATAACGATCGATGGGCTGCTGCTCCAAGTGAATTTCAAG 1680
Qy	1681 TGTCTTATCAAAAGAGGAGATAGCAATTTACCAAGTGGTGAATTTGGGAAGTCTTGGCCGGCAT 1740
Db	1681 TGTCTTATCAAAAGAGGAGATAGCAATTTACCAAGTGGTGAATTTGGGAAGTCTTGGCCGGCAT 1740
Qy	1741 GGAATCTAATATCCAAAGTGTGAGAGTCAAGAGGCTGATATTTTGAAGGACCAACAGAC 1800
Db	1741 GGAATCTAATATCCAAAGTGTGAGAGTCAAGAGGCTGATATTTTGAAGGACCAACAGAC 1800
Qy	1801 TCCCTTTAGAGCATCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db	1801 TCCCTTTAGAGCATCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Qy	1861 CTGACTGACCGTGGCTACTCACAATTTCTGCTGATCAGTACGACTGCTGATCTTTTATA 1920

Db	1861	CTGACTGACCGTGGCTACTCACATCTTTGCTGCATCAGTCAGCACTGTGACTTCTTTATA	1920
Qy	1921	AGTAACTAGTAGTAACACAGAGAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCTCT	1980
Db	1921	AGTAACTAGTAGTAACACAGAGAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCTCT	1980
Qy	1981	GAAGATGACCCCAACTTGCAGAACAAAGGAATTTTGGCCACCAATTTTGGATTGAGAGTCA	2040
Db	1981	GAAGATGACCCCAACTTGCAGAACAAAGGAATTTTGGCCACCAATTTTGGATTGAGAGTCA	2040
Qy	2041	CCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTCTGGAATTTACA	2100
Db	2041	CCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTCTGGAATTTACA	2100
Qy	2101	TTGTATGGAGTCTCTCAAGCCCTCATGATCTACAGCTTGGAAAGAAATATCTACTGTG	2160
Db	2101	TTGTATGGAGTCTCTCAAGCCCTCATGATCTACAGCTTGGAAAGAAATATCTACTGTG	2160
Qy	2161	CTGTTTCATATATGTGG--TCTCAGGTGAGTGTGGTAATATCGTTTAAAGGAGTCA	2218
Db	2161	CTGTTTCATATATGTGGTCTCTCAGGTGAGTGTGGTAATATCGTTTAAAGGAGTCA	2220
Qy	2219	AGTATTTCCGCTTGATACCTAGCTCTCTAGGTTATGTGTTGTAGTGATAGACAACA	2278
Db	2221	AGTATTTCCGCTTGATACCTAGCTCTCTAGGTTATGTGTTGTAGTGATAGACAACA	2280
Qy	2279	GGGATCTCTGTCACGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAA	2338
Db	2281	GGGATCTCTGTCACGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAA	2340
Qy	2339	TAGAAATGACATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATG	2398
Db	2341	TAGAAATGACATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATG	2400
Qy	2399	ACTTAGATCTGTGGGCATCCAGCGTGTCTCTATGGAGGATACCTCTCCCTGATGCAAT	2458
Db	2401	ACTTAGATCTGTGGGCATCCAGCGTGTCTCTATGGAGGATACCTCTCCCTGATGCAAT	2460
Qy	2459	TAATGACAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCACTCTCTGGA	2518
Db	2461	TAATGACAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCACTCTCTGGA	2520
Qy	2519	TCTTCTATGATACAGATACACGAAAGCTTATATGGTCAACCTGACCGNAGTAAACAGG	2578
Db	2521	TCTTCTATGATACAGATACACGAAAGCTTATATGGTCAACCTGACCGNAGTAAACAGG	2580
Qy	2579	GCTATTACTTAGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTT	2638
Db	2581	GCTATTACTTAGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTT	2640
Qy	2639	TACTGCTCTTACATGTTTCTGATGGAATGTCATTTGCAATACCAATATATAC	2698
Db	2641	TACTGCTCTTACATGTTTCTGATGGAATGTCATTTGCAATACCAATATATAC	2700
Qy	2699	TGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGAC	2758
Db	2701	TGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGAC	2760
Qy	2759	ACAGCATAGAGTTCTGATCGGAGAACATTAATGAATGCACTTTTGCACATCTTC	2818
Db	2761	ACAGCATAGAGTTCTGATCGGAGAACATTAATGAATGCACTTTTGCACATCTTC	2820
Qy	2819	AGAAACCTTGATACGATATGCTCTTAAAGTGATATAATTTGACCTGTGTA	2878
Db	2821	AGAAACCTTGATACGATATGCTCTTAAAGTGATATAATTTGACCTGTGTA	2880
Qy	2879	ACTCTCTGTATACCTGGCTATTAAACCAATGAGGAGTTTAATCAACAGAAACACA	2938
Db	2881	ACTCTCTGTATACCTGGCTATTAAACCAATGAGGAGTTTAATCAACAGAAACACA	2940
Qy	2939	GAAATTGATCATCATTTTGTATACCTGCCATGTAAATCTCTCTGAAATTAATGTGG	2998
Db	2941	GAAATTGATCATCATTTTGTATACCTGCCATGTAAATCTCTCTGAAATTAATGTGG	3000
Qy	2999	TGCCATGACGGGTCTACGGTTTGTGTAGTAATCTTAATACCTTAACCCCAATGCTCAA	3058
Db	3001	TGCCATGACGGGTCTACGGTTTGTGTAGTAATCTTAATACCTTAACCCCAATGCTCAA	3060
Qy	3059	AATCAAAATGATACATATCTCTGAGAGCCAGCAATACCAATGAATTAATAAAAAA	3118
Db	3061	AATCAAAATGATACATATCTCTGAGAGCCAGCAATACCAATGAATTAATAAAAAA	3120
Qy	3119	AA 3120	
Db	3121	AA 3122	
RESULT 7			
AX608735			
LOCUS			
DEFINITION			
AX608735			
ACCESSION			
AX608735.1			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE			
1			
Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.			
Novel serine protease genes related to appiv			
Patent: WO 0231134-A 12 18-APR-2002;			
Ferring BV (NL)			
FEATURES			
Location/Qualifiers			
source			
1..4829			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
ORIGIN			
Query Match			
Best Local Similarity 92.6%; Score 2889; DB 2; Length 4829;			
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;			
Qy	1	AAGTGCTAAAGCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGG	60
Db	1	AAGTGCTAAAGCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGG	60
Qy	61	CGTTTCGCGCGCTGGGTGTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	120
Db	61	CGTTTCGCGCGCTGGGTGTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	120
Qy	121	GAGTGGAGGCG	180
Db	121	GAGTGGAGGCG	180
Qy	181	TCCGGGCG	240
Db	181	TCCGGGCG	240
Qy	241	CTGGGTGTTGAGATATTTGAAACTCGGACCTGTGAGGAGATATTTGAATCAAGATCGG	300
Db	241	CTGGGTGTTGAGATATTTGAAACTCGGACCTGTGAGGAGATATTTGAATCAAGATCGG	300
Qy	301	CCTAATTCGAGCGCTTTTATGTTGAGCGGTATTTCTTGGATGAGCTTAAAAAGCTGCTT	360
Db	301	CCTAATTCGAGCGCTTTTATGTTGAGCGGTATTTCTTGGATGAGCTTAAAAAGCTGCTT	360
Qy	361	GGCGATACAGAAATATCATGGCTATCATGATGCTAAGGACCAACATGATTTTCTGTTT	420
Db	361	GGCGATACAGAAATATCATGGCTATCATGATGCTAAGGACCAACATGATTTTCTGTTT	420
Qy	421	GTGAGAGGATGATCCAGATGGACCTCATTTACAGAGAAATCTATTACCTTGGCATGTCT	480
Db	421	GTGAGAGGATGATCCAGATGGACCTCATTTACAGAGAAATCTATTACCTTGGCATGTCT	480

QY	481	GGTGGAAACAGAGAAATACACTGTTTATCTGAAATCCCAAACTCATATAGAGCA	540	1561	TTTATTTTGGCTCTGAAATGATAACAGGTTTCGGTCAATTTATATACAAATTTACATCTATT	1620
Db	481	GGTGGAAACAGAGAAATACACTGTTTATCTGAAATCCCAAACTCATATAGAGCA	540	1561	TTTATTTTGGCTCTGAAATGATAACAGGTTTCGGTCAATTTATATACAAATTTACATCTATT	1620
QY	541	GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGATCTTTTTCAGGCAACACTGGACTAT	600	1621	TTAAAGGAAAGCAAAATATAAACGATCCAGTGGTGGGCTGCTCCTCAAGTGATTTCAAG	1680
Db	541	GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGATCTTTTTCAGGCAACACTGGACTAT	600	1621	TTAAAGGAAAGCAAAATATAAACGATCCAGTGGTGGGCTGCTCCTCAAGTGATTTCAAG	1680
QY	601	GGATGATTTCTCGAGAGAGAACTATTAAGAGAAAGAAACGATTTGGAACAGTCGGA	660	1681	TGTCCTTATCAAAGAGAGATAGCAATTTACAGTGGTGAATCGGAAGTTCTTTGGCCGGCAT	1740
Db	601	GGATGATTTCTCGAGAGAGAACTATTAAGAGAAAGAAACGATTTGGAACAGTCGGA	660	1681	TGTCCTTATCAAAGAGAGATAGCAATTTACAGTGGTGAATCGGAAGTTCTTTGGCCGGCAT	1740
QY	661	ATTGCTTCTTACGATTTATCAACAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGA	720	1741	GGATCTAATATCCAAGTTCGATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGAC	1800
Db	661	ATTGCTTCTTACGATTTATCAACAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGA	720	1741	GGATCTAATATCCAAGTTCGATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGAC	1800
QY	721	ATTTATCACGTAAGAGTGGAGGCGCACAGGATTTAGCAACACCTTTAAGGCCCAAT	780	1801	TCCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGCTTAAGTTCCTGGAGAGGTGACAGG	1860
Db	721	ATTTATCACGTAAGAGTGGAGGCGCACAGGATTTAGCAACACCTTTAAGGCCCAAT	780	1801	TCCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGCTTAAGTTCCTGGAGAGGTGACAGG	1860
QY	781	CTAGTGAACACTAGTTTGTCCCAACATACGATGGAATCCAAATTTATGCCCGCTGATCCA	840	1861	CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATA	1920
Db	781	CTAGTGAACACTAGTTTGTCCCAACATACGATGGAATCCAAATTTATGCCCGCTGATCCA	840	1861	CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATA	1920
QY	841	GACTGGATTCCTTTTATACATAGCAACGATATTTGGATCTTAAACATCGTAAACAGAA	900	1921	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCT	1980
Db	841	GACTGGATTCCTTTTATACATAGCAACGATATTTGGATCTTAAACATCGTAAACAGAA	900	1921	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCT	1980
QY	901	GAAAGGAGACTCTATTATGTCACATGAGCTAGGCCAACATGGAAGAAGATGCCAGATCA	960	1981	GAAATGACCCAACTTGCACAAACAAAGGAAATTTTGGGCCACCATTTTGGATTCAGAGGT	2040
Db	901	GAAAGGAGACTCTATTATGTCACATGAGCTAGGCCAACATGGAAGAAGATGCCAGATCA	960	1981	GAAATGACCCAACTTGCACAAACAAAGGAAATTTTGGGCCACCATTTTGGATTCAGAGGT	2040
QY	961	GCTGGAGTCGATCTTGTCTTCCAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020	2041	CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTACA	2100
Db	961	GCTGGAGTCGATCTTGTCTTCCAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020	2041	CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTACA	2100
QY	1021	TGTCCTAAGAGTGAACAACTCCAGTGGTGGTAAATTTCTTGAATTTCTATAGAAGAA	1080	2101	TTGTATGGGATGCTCTACAAAGCTCATGATCTACAGCTCATGATCTGAAAGAAATATCTACTGTG	2160
Db	1021	TGTCCTAAGAGTGAACAACTCCAGTGGTGGTAAATTTCTTGAATTTCTATAGAAGAA	1080	2101	TTGTATGGGATGCTCTACAAAGCTCATGATCTACAGCTCATGATCTGAAAGAAATATCTACTGTG	2160
QY	1081	AATGATGAATCTGAGGTGGAATTTATGTTTACATCCCTTATGTTGGAAACAGGAGG	1140	2161	CTGTTTCATATATATGTTGG--TCCTCAGGTGAGTGGTGGTGAATATCCGTTTAAAGGAGTCA	2218
Db	1081	AATGATGAATCTGAGGTGGAATTTATGTTTACATCCCTTATGTTGGAAACAGGAGG	1140	2161	CTGTTTCATATATATGTTGG--TCCTCAGGTGAGTGGTGGTGAATATCCGTTTAAAGGAGTCA	2218
QY	1141	GCAGATTCATTCGGTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAGATG	1200	2219	AGTATTTTCGGTTCGAATACCTAGCTCTCTAGGTATGTTGGTGTAGTGATAGACAA	2278
Db	1141	GCAGATTCATTCGGTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAGATG	1200	2219	AGTATTTTCGGTTCGAATACCTAGCTCTCTAGGTATGTTGGTGTAGTGATAGACAA	2278
QY	1201	TCAGAAATAATGATGCTGGAAGGAGGATCATAGATGTCATAGATAAGAACTAAT	1260	2279	GGGGATCCTGTCCACCGAGGCTTAAATTTGAAGGCGCTTAAATATATAAATGGGTCAA	2338
Db	1201	TCAGAAATAATGATGCTGGAAGGAGGATCATAGATGTCATAGATAAGAACTAAT	1260	2279	GGGGATCCTGTCCACCGAGGCTTAAATTTGAAGGCGCTTAAATATATAAATGGGTCAA	2338
QY	1261	CAACCTTTTGAGATTTATTTGAAGGAGTGAATATATGCGAGAGCTGGATGGCTCT	1320	2339	TAGAAATTTGACGATCAGGTGGAAGGATCCCAATATCTAGCTTCCGATATGATTTTCAATG	2398
Db	1261	CAACCTTTTGAGATTTATTTGAAGGAGTGAATATATGCGAGAGCTGGATGGCTCT	1320	2339	TAGAAATTTGACGATCAGGTGGAAGGATCCCAATATCTAGCTTCCGATATGATTTTCAATG	2398
QY	1321	GAGGAAATATGCTTGGTCCATCTAGATGCTCCAGAGCTCCCTACAGATGATG	1380	2401	ACTTAGATCGTGGGCGATCCACGCTGCTCTATGAGGAGTACCTCTCCCTGATGGCAT	2458
Db	1321	GAGGAAATATGCTTGGTCCATCTAGATGCTCCAGAGCTCCCTACAGATGATG	1380	2401	ACTTAGATCGTGGGCGATCCACGCTGCTCTATGAGGAGTACCTCTCCCTGATGGCAT	2458
QY	1381	TTGATCTACCTGAAATTTATTTTCCAGTAGAAGATGATGTTATGGAAGGCGAGACTC	1440	2459	TAATGAGAGGTGAGATATCTTACGGGTTGCTATTTGCTGGGGCCCCAGCTCTCTGTGGA	2518
Db	1381	TTGATCTACCTGAAATTTATTTTCCAGTAGAAGATGATGTTATGGAAGGCGAGACTC	1440	2459	TAATGAGAGGTGAGATATCTTACGGGTTGCTATTTGCTGGGGCCCCAGCTCTCTGTGGA	2518
QY	1441	ATTGAGTCAGTCCCTGATTCGTGAGCGCACTAATATCTATGAGAAACAGACATC	1500	2519	TCTTCTATGATACAGGATACAGGAACTTATGGGTGCTGAGGAGTACCTCTCCCTGATGGAG	2578
Db	1441	ATTGAGTCAGTCCCTGATTCGTGAGCGCACTAATATCTATGAGAAACAGACATC	1500	2519	TCTTCTATGATACAGGATACAGGAACTTATGGGTGCTGAGGAGTACCTCTCCCTGATGGAG	2578
QY	1501	TGATTAATATCCATGACATCTTCTATGTTTCCCAAGTACAGAGGAGAAATTTGAG	1560	2579	GCTATTACTTAGGATCTGTGGCGATGCAAGCAGAAAGTTCCTCTCTGAAACCAATCGTT	2638
Db	1501	TGATTAATATCCATGACATCTTCTATGTTTCCCAAGTACAGAGGAGAAATTTGAG	1560	2579	GCTATTACTTAGGATCTGTGGCGATGCAAGCAGAAAGTTCCTCTCTGAAACCAATCGTT	2638
QY	1561	TTTATTTTGGCTCTGAAATGATAACAGGTTTCGGTCAATTTATATACAAATTTACATCTATT	1620	2639	TACTGCTCTTACATGTTTCTGGATGAGATGTCCATTTTGGACATACAGATATATATAC	2698
Db	1561	TTTATTTTGGCTCTGAAATGATAACAGGTTTCGGTCAATTTATATACAAATTTACATCTATT	1620	2639	TACTGCTCTTACATGTTTCTGGATGAGATGTCCATTTTGGACATACAGATATATATAC	2698

Db 976 TCTCCAGAGAAATTTGATAGATATCTGGCTATTGGTGTCCAAAGCTGAAACAAC 1035
Qy 1041 TCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGA 1100
Db 1036 TCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGA 1095
Qy 1101 AATTATTCACTGATCATCCCTATGTTGGAAACAAGAGGGGCGAGATTCAATTCGGTTATCC 1160
Db 1096 AATTATTCACTGATCATCCCTATGTTGGAAACAAGAGGGGCGAGATTCAATTCGGTTATCC 1155
Qy 1161 TAAACAGGTTACAGAAATCTAAAGTCACTTTTAAGATGTCAGAAAATGATGATGATGC 1220
Db 1156 TAAACAGGTTACAGAAATCTAAAGTCACTTTTAAGATGTCAGAAAATGATGATGATGC 1215
Qy 1221 TGAAGGAGGATCATAGATGTCATAGATAAGGAACCTAAATTTCAACCTTTTGGAGATTTCTATT 1280
Db 1216 TGAAGGAGGATCATAGATGTCATAGATAAGGAACCTAAATTTCAACCTTTTGGAGATTTCTATT 1275
Qy 1281 TGAAGGAGGTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGGGAAAATATGCTTGGTC 1340
Db 1276 TGAAGGAGGTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGGGAAAATATGCTTGGTC 1335
Qy 1341 CATCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTAAT 1400
Db 1336 CATCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTAAT 1395
Qy 1401 TATCCCACTAGAGATGATGTTATGGAAGGCGAGAGACTCATTTAGTCACTGCTGATTC 1460
Db 1396 TATCCCACTAGAGATGATGTTATGGAAGGCGAGAGACTCATTTAGTCACTGCTGATTC 1455
Qy 1461 TGTGACGCGCACTAATTAATCTATGAAGAAAACAAGACATCTGGATTAATATCATGACAT 1520
Db 1456 TGTGACGCGCACTAATTAATCTATGAAGAAAACAAGACATCTGGATTAATATCATGACAT 1515
Qy 1521 CTTTCACTGTTTCCCAAGCTCACGAAGGAGAAATTTAGTTTATTTTGGCTCTGAAATG 1580
Db 1516 CTTTCACTGTTTCCCAAGCTCACGAAGGAGAAATTTAGTTTATTTTGGCTCTGAAATG 1575
Qy 1581 CAAAAACAGGTTTCCGTCAATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATAA 1640
Db 1576 CAAAACAGGTTTCCGTCAATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATAA 1635
Qy 1641 ACGATCCAGTGGTGGCTGCTGCTCCAGTGAATTTCAAGTGTCTCTATCAAGAGAGAT 1700
Db 1636 ACGATCCAGTGGTGGCTGCTGCTCCAGTGAATTTCAAGTGTCTCTATCAAGAGAGAT 1695
Qy 1701 AGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCGGCATGGATCTAATATCCAAGTTGA 1760
Db 1696 AGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCGGCATGGATCTAATATCCAAGTTGA 1755
Qy 1761 TGAAGTCAAGAGGCTGGTATATTTTGAAGGCAACAAAGATCTCCCTTTTGAAGCATCACCT 1820
Db 1756 TGAAGTCAAGAGGCTGGTATATTTTGAAGGCAACAAAGATCTCCCTTTTGAAGCATCACCT 1815
Qy 1821 GTACCTGATGATGATGATTAATCTTGGAGGTCACAAAGCTGACCTGACCGTGGCTACTC 1880
Db 1816 GTACCTGATGATGATGATTAATCTTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTC 1875
Qy 1881 ACATTTCTGCTGATCAGTCAGCACCTGTGACTCTCTTTATAGTAAGTATAGTAACAGAA 1940
Db 1876 ACATTTCTGCTGATCAGTCAGCACCTGTGACTCTCTTTATAGTAAGTATAGTAACAGAA 1935
Qy 1941 GAATCCACACTGTGTGCTCTTACAGCTATCAAGTCTCAAGTCAAGTCAAGTCAAGTCAAG 2000
Db 1936 GAATCCACACTGTGTGCTCTTACAGCTATCAAGTCTCAAGTCAAGTCAAGTCAAGTCAAG 1995
Qy 2001 AACAAAGGAATTTTGGGCGACCATTTTGGATTTCAAGAGGCTCTCTTCTGACTATACTCC 2060
Db 1996 AACAAAGGAATTTTGGGCGACCATTTTGGATTTCAAGAGGCTCTCTTCTGACTATACTCC 2055
Qy 2061 TCCAGAAATTTTCTCTTTTGAAGTATCTAGTGAATTTATGATGGGATGCTCTACAA 2120

Db 2056 TCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAA 2115
Qy 2121 GCCTCATGATCTTACAGCCTCGAAAGAAATATCTTACTGTCTGTTCTATATATGTTGGTCC 2180
Db 2116 GCCTCATGATCTTACAGCCTCGAAAGAAATATCTTACTGTCTGTTCTATATATGTTGGTCC 2175
Qy 2181 TCAGGTGCAATTTGTTGTTAAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCT 2240
Db 2176 TCAGGTGCAATTTGTTGTTAAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCT 2235
Qy 2241 AGCCTCTCTAGGTTATGTTGTTAGTAGATAGACAAACAGGGGATCTCTGTCACCGAGGCT 2300
Db 2236 AGCCTCTCTAGGTTATGTTGTTAGTAGATAGACAAACAGGGGATCTCTGTCACCGAGGCT 2295
Qy 2301 TAAATTTGAAGCGGCTTTTAAATATAAAATGAGGTTCAAATAGAAATTCAGCATGAGTGA 2360
Db 2296 TAAATTTGAAGCGGCTTTTAAATATAAAATGAGGTTCAAATAGAAATTCAGCATGAGTGA 2355
Qy 2361 AGGATCCCAATATCTAGCTTCTCGATATGATTTCAATTCATAGATCGTGGGATCCCA 2420
Db 2356 AGGATCCCAATATCTAGCTTCTCGATATGATTTCAATTCATAGATCGTGGGATCCCA 2415
Qy 2421 CGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTAATGACAGAGGTCAGATATCTT 2480
Db 2416 CGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTAATGACAGAGGTCAGATATCTT 2475
Qy 2481 CAGGGTTGCTATTTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATCAC 2540
Db 2476 CAGGGTTGCTATTTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATCAC 2535
Qy 2541 GGAACCTTATATGAGGTCACCTGACAGAAATGAACAGGGCTATTTACTTAGGATCTGGC 2600
Db 2536 GGAACCTTATATGAGGTCACCTGACAGAAATGAACAGGGCTATTTACTTAGGATCTGGC 2595
Qy 2601 CATGCAAGCAGAAAAGTTCCCTCTGAAACCAATTCGTTTACTGCTCTTATCATGTTTCT 2660
Db 2596 CATGCAAGCAGAAAAGTTCCCTCTGAAACCAATTCGTTTACTGCTCTTATCATGTTTCT 2655
Qy 2661 GGAATGATGTCCTTATTTGACATACAGTATATTAATGAGTGTGTTTGTAGTGGGCTGG 2720
Db 2656 GGAATGATGTCCTTATTTGACATACAGTATATTAATGAGTGTGTTTGTAGTGGGCTGG 2715
Qy 2721 AAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAAGAGTTCTGATC 2780
Db 2716 AAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAAGAGTTCTGATC 2775
Qy 2781 GGAGAACATATATGATGCTATCTTTTGACATACCTTCAAGAAAACCTTGGATCACGAT 2840
Db 2776 GGAGAACATATATGATGCTATCTTTTGACATACCTTCAAGAAAACCTTGGATCACGAT 2835
Qy 2841 TGCTGCTCTAAGAGTATATTTTGGACCTGTGTAGAACTCTCTGGTATACACCTGCTA 2900
Db 2836 TGCTGCTCTAAGAGTATATTTTGGACCTGTGTAGAACTCTCTGGTATACACCTGCTA 2895
Qy 2901 TTTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATGATCATCATTTGAT 2960
Db 2896 TTTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATGATCATCATTTGAT 2955
Qy 2961 ACCTGCTATGATGATCTTACTCTGAAATTAATGCTGGTGCATGCA 3007
Db 2956 ACCTGCTATGATGATCTTACTCTGAAATTAATGCTGGTGCATGCA 3002

RESULT 9

AX354795
LOCUS AX354795 2649 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 3 from Patent WO0179473.
ACCESSION AX354795
VERSION AX354795.1 GI:18619528
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS
Meyers, R.A. and Williamson, M.
21953, a human prolly oligopeptidase family member and uses thereof
TITLE
Patent: WO 0179473-A 3 25-OCI-2001;
JOURNAL
Millennium Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
source
1..2649
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 81.6%; Score 2547; DB 2; Length 2649;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2647; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
214 ATGGCAGCAGCAATGGAAACAGAAACAGCTGGGTGTGGAGATATTTGAAATCTGGGACTGT 273
Db 1 ATGGCAGCAGCAATGGAAACAGAAACAGCTGGGTGTGGAGATATTTGAAATCTGGGACTGT 60
274 GAGGAGATATTTGAATCAAGGATCGGCCTAAATTTGGAGCCTTTTATGTGAGCGGTAT 333
Db 61 GAGGAGATATTTGAATCAAGGATCGGCCTAAATTTGGAGCCTTTTATGTGAGCGGTAT 120
334 TCCTGGAGTCAGCTTAAAGAGCTGCTTGGCCGATACAGATCGGCCTTTTATGTGAGCGGTAT 393
Db 121 TCCTGGAGTCAGCTTAAAGAGCTGCTTGGCCGATACAGATCGGCCTTTTATGTGAGCGGTAT 180
394 GCTAAGGCACCAATGATTTTATGTTTGAAGAGCAATGATCCAGATGGACCTCATCTCA 453
Db 181 GCTAAGGCACCAATGATTTTATGTTTGAAGAGCAATGATCCAGATGGACCTCATCTCA 240
454 GACAGATCTATTACCTTGGCCATGCTGTTGAGAGCAAGAAATACACTGTTTATTTCT 513
Db 241 GACAGATCTATTACCTTGGCCATGCTGTTGAGAGCAAGAAATACACTGTTTATTTCT 300
514 GAATTTCCCAAACTATCAATAGCAGCAGCTTAAATGCTCTCTTGGAGCCTCTTTTG 573
Db 301 GAATTTCCCAAACTATCAATAGCAGCAGCTTAAATGCTCTCTTGGAGCCTCTTTTG 360
574 GATCTTTTCAGCGCAGCTGGAATGGAATTTTCTCGAGAGAGAGAACTATTAAAGA 633
Db 361 GATCTTTTCAGCGCAGCTGGAATGGAATTTTCTCGAGAGAGAGAACTATTAAAGA 420
634 GAAG 693
Db 421 GAAG 480
694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATCACTGTAAGAGATGGAGGAGAGAGAG 753
Db 481 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATCACTGTAAGAGATGGAGGAGAGAGAG 540
754 TTTACGCAACAACTTTAAGGCCAATCTAGTGGAAACTAGTGTGTCACCAATACGGATG 813
Db 541 TTTACGCAACAACTTTAAGGCCAATCTAGTGGAAACTAGTGTGTCACCAATACGGATG 600
814 GATCCAAATTTATGCCCCTGATCCAGACTGGAATTTGCTTTTATACATAGCAAGATTT 873
Db 601 GATCCAAATTTATGCCCCTGATCCAGACTGGAATTTGCTTTTATACATAGCAAGATTT 660
874 TGGATATCTAACAATCGTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 933
Db 661 TGGATATCTAACAATCGTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
934 GCCAATGAG 993
Db 721 GCCAATGAG 780
994 TTTGATAGATATTTCTGGCTATTTGGTGTCCAAAGCTGAAACCTCCAGTGGTGT 1053
Db 781 TTTGATAGATATTTCTGGCTATTTGGTGTCCAAAGCTGAAACCTCCAGTGGTGT 840

1054 AAAATCTTTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
Db 841 AAAATCTTTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 900
1114 ACATCCCTTATGTTGGAAACAGAGGAGGAGATTCATTCCTGTTATCTTAAACAGGTACA 1173
Db 901 ACATCCCTTATGTTGGAAACAGAGGAGGAGATTCATTCCTGTTATCTTAAACAGGTACA 960
1174 GCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATCTGAGAGAGATC 1233
Db 961 GCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATCTGAGAGAGATC 1020
1234 ATAGATGTCATAGATAAGAACTAAATTCACCTTTTGAAGATTTTATTTGAAGAGTTGAA 1293
Db 1021 ATAGATGTCATAGATAAGAACTAAATTCACCTTTTGAAGATTTTATTTGAAGAGTTGAA 1080
1294 TATATTTGCGAGAGCTGGATGGAATCTTGGAGGAGAAATATCTTGGTCCATCTTAGAT 1353
Db 1081 TATATTTGCGAGAGCTGGATGGAATCTTGGAGGAGAAATATCTTGGTCCATCTTAGAT 1140
1354 CGCTCCGAGACTCGGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1413
Db 1141 CGCTCCGAGACTCGGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1200
1414 GATGATGTTATGGAAGAGGAGAGACTCAITGAGTCACTGCTGATCTGTGAGCGCACTA 1473
Db 1201 GATGATGTTATGGAAGAGGAGAGACTCAITGAGTCACTGCTGATCTGTGAGCGCACTA 1260
1474 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTT 1533
Db 1261 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTT 1320
1534 CCCCAAGTCAAGAGAGGAAATTTAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1593
Db 1321 CCCCAAGTCAAGAGAGGAAATTTAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1380
1594 CGTCATTTATACAAATTTACATCTATTTTAAAGAGAGCAATATAAAGATCCAGTGGT 1653
Db 1381 CGTCATTTATACAAATTTACATCTATTTTAAAGAGAGCAATATAAAGATCCAGTGGT 1440
1654 GGGCTGCTGCTCCCAAGTGAATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATACCA 1713
Db 1441 GGGCTGCTGCTCCCAAGTGAATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATACCA 1500
1714 GGTGAATGGGAAGTTCTTGGCCGAGTGGATCTTATATCCAAAGTTGATGAGTCAAGAG 1773
Db 1501 GGTGAATGGGAAGTTCTTGGCCGAGTGGATCTTATATCCAAAGTTGATGAGTCAAGAG 1560
1774 CTGCTATATTTTGAAGGACCAAGAGCTCGCTTTAGAGCATCAGCTGTACGTAGTCA 1833
Db 1561 CTGCTATATTTTGAAGGACCAAGAGCTCGCTTTAGAGCATCAGCTGTACGTAGTCA 1620
1834 TACGTAAATCTGGAGAGTGACAGGCTGACCTGACCGTGGCTACTCACTTTCTTGTGTC 1893
Db 1621 TACGTAAATCTGGAGAGTGACAGGCTGACCTGACCGTGGCTACTCACTTTCTTGTGTC 1680
1894 ATCAGTCAGCACTGTGACTCTTTTATAGTAAAGTATAGTAAACAGAGAAATCCACACT 1953
Db 1681 ATCAGTCAGCACTGTGACTCTTTTATAGTAAAGTATAGTAAACAGAGAAATCCACACT 1740
1954 GTGTCCCTTTTCAAGCTATCAAGTCTCTGAAAGTATGACCAACTTGCAGAAACAGAGAA 2013
Db 1741 GTGTCCCTTTTCAAGCTATCAAGTCTCTGAAAGTATGACCAACTTGCAGAAACAGAGAA 1800
2014 TGGGCCACCAATTTTGGATCTCTCTTCTCTGACTATCTCTCTCCAGAAATTTTC 2073
Db 1801 TGGGCCACCAATTTTGGATCTCTCTTCTCTGACTATCTCTCTCCAGAAATTTTC 1860
2074 TCTTTTGAAGAGTACTACTGGAATTTACATTTGATGGATGCTCTCAAGGCTCTCATGATCTA 2133
Db 1861 TCTTTTGAAGAGTACTACTGGAATTTACATTTGATGGATGCTCTCAAGGCTCTCATGATCTA 1920

Submitted (04-NOV-2002) Ferring Research Institute, 3550 General Atomics Ct., San Diego, CA 92121, USA		JOURNAL	
Location/Qualifiers		FEATURES	
1..2656		source	
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/chromosome="15"			
/map="15q22.1-q22.2"			
1..2656		gene	
/gene="DPRP1"			
8..2656		CDS	
/gene="DPRP1"			
/codon_start=1			
/product="dipeptidyl peptidase IV-related protein-1"			
/protein_id="AA017261.1"			
/db_xref="GI:27549550"			
/translation="MAAMETEQLGVIFETADCEENIESQDRPKLEPFYVERYSWQ LKKJLADTRKHYGNMAKAPHDPMFVRKNDPDPGPHSDRIYYLAMSGENRENTLYSEI PKTINRAAVLMSWKPLDLDFQATLDYGMYSREBELLRKRKRGITGVGIAIDYHQSGG TFLFOAGSGIVHVGQGGPQOQLPRLNVLVETSCPNIRMDPKLCPADPDWIAPIHSN DIWISNIVTRERRLTYYVHNELANNEEDARSAGVATFVLOEEDFYSGYVWCPEAKETT PSQKILRLIYEENDESEVEIIVHTSPMLETRRADSPRYKPTGTANPKVTFKMSSEIMI DAEGRLIDVIDKELLQPELIFEGVEVILARAGWTPGKYAWSILLDRSQRLQIVLIS PELFIPVEDDVMERQLIESVPDSVTLPIIYEETDIIWINIHDI FHVFPQSHREELTF IFASECKTGFRHLVYKITISILKESKYSSGGLPAPSPDFKCPIKEEIAITSGEWEVLGR HGSNIQDVSRRLVYFEGTKDSPLEHLYVVSYPGEVTRLDRGVSCHSCSISQHCD FTISKYSNKNPHCVSLVKLSPPEDDPTCKTFEWFATLDSAGPLPDYTPPEIFSPES TTGFTLTKMLXPHLDQPKKYPTVLFTYGGPQVQLVNNRPFKGVYFRLANTLASLGVV VVVDNRGSGHRLKPEGAFKFKWGOIIEIDDOVSGLOYLASRYDFIDLDRVGIHGSYV GQYLSJLMALUMQRSDIFRVAIAGAPVTLMIPYDTGYTRYRMGHSPQNSQGYVVGSAVMO AEKFPSEPNRLLHFLDENVFAHTSILLSFLVRAGKPYDLQIYPQERHSIRVPES GEHYELHLHLYLQENLGSRIALKVI"			
		ORIGIN	
		Query Match 80.0%; Score 2495; DB 5; Length 2656;	
		Best Local Similarity 99.3%; Pred. No. 0;	
		Matches 2645; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
QY	213	CATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAAACTGGGACTG	272
Db	7	CATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAAACTGGGACTG	66
QY	273	TGAGAGAAATATGAATCAGGATCGGCTAAATGGAGCCCTTTTATGTTGACGGTA	332
Db	67	TGAGAGAAATATGAATCAGGATCGGCTAAATGGAGCCCTTTTATGTTGACGGTA	126
QY	333	TTCTCTGGAGTCAGCTTAAAGCTGCTTGGCGATACAGAAATATCATGGCTACATGAT	392
Db	127	TTCTCTGGAGTCAGCTTAAAGCTGCTTGGCGATACAGAAATATCATGGCTACATGAT	186
QY	393	GGCTAAGGCACACATGATTTTCATGTTTGAAGAGATGATCCAGATGGACCTCATTC	452
Db	187	GGCTAAGGCACACATGATTTTCATGTTTGAAGAGAGATGATCCAGATGGACCTCATTC	246
QY	453	AGACAGAAATCTATTACCTTGCCATGCTGCTGGTGAACACAGAGAAATACACTGTTTATTC	512
Db	247	AGACAGAAATCTATTACCTTGCCATGCTGCTGGTGAACACAGAGAAATACACTGTTTATTC	306
QY	513	TGAATATCCCAAACTATCAATAGACAGCAGTCTTAATGCTCTCTTGGAGGCTCTTTT	572
Db	307	TGAATATCCCAAACTATCAATAGACAGCAGTCTTAATGCTCTCTTGGAGGCTCTTTT	366
QY	573	GGATCTTTTTCAGCAACACTGGACTATGGATGCTATCTTCGAGAGAGAACTATTAAAG	632
Db	367	GGATCTTTTTCAGCAACACTGGACTATGGATGCTATCTTCGAGAGAGAACTATTAAAG	426
QY	633	AGAAAGAAACCGATTGGAAACAGTCGGAATGCTTTTACGATTATCACCAAGGAAGTGG	692
Db	427	AGAAAGAAACCGATTGGAAACAGTCGGAATGCTTTTACGATTATCACCAAGGAAGTGG	486
QY	693	AACATTTCTGTTTCAAGCGCGGTAGTGGAAATTTATCATGTAAGAGATCGAGGCCACAGG	752

Db	487	AACATTTCTGTTTCAAGCCGGTAGTGGAATTTATCAGGTAAGAGATGGAGGCCACAAGG	546
Qy	753	ATTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAACTAGTGTGCCAATACCGAT	812
Db	547	ATTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAACTAGTGTGCCAATACCGAT	606
Qy	813	GGATCCAAAATTTATGCCCGCTGATCCAGACTCGGATTTCTTTATACATAGCAAGATAT	872
Db	607	GGATCCAAAATTTATGCCCGCTGATCCAGACTCGGATTTCTTTATACATAGCAAGATAT	666
Qy	873	TTGGATATCTAACATCTTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	932
Db	667	TTGGATATCTAACATCTTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	726
Qy	933	AGCCAACTGAAGAAGATGCCAGATCAGCTGAGTTCGCTACCTTTGTTCTCCAAAGA	992
Db	727	AGCCAACTGAAGAAGATGCCAGATCAGCTGAGTTCGCTACCTTTGTTCTCCAAAGA	786
Qy	993	ATTTGATAGATATCTGGCTATTTGGTGGTCCAAAGCTGAAGCAAACTCCAGTGGTGG	1052
Db	787	ATTTGATAGATATCTGGCTATTTGGTGGTCCAAAGCTGAAGCAAACTCCAGTGGTGG	846
Qy	1053	TAAAATCTTAGAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATATATCATGT	1112
Db	847	TAAAATCTTAGAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATATATCATGT	906
Qy	1113	TACATCCCTATGTTGAAACAAGGAGGCGAGATTCATTCGGTATATCTTAATAACAGGTAC	1172
Db	907	TACATCCCTATGTTGAAACAAGGAGGCGAGATTCATTCGGTATATCTTAATAACAGGTAC	966
Qy	1173	AGCAATCTCTAAAGTCACCTTTTAAAGATGTCAAGAAATATGATGATGCTGAAGGAAGAT	1232
Db	967	AGCAATCTCTAAAGTCACCTTTTAAAGATGTCAAGAAATATGATGATGCTGAAGGAAGAT	1026
Qy	1233	CATAGATGTCATAGATAAGGAACCTAATTCACCTTTTTCAGATCTCTATTTGAAGGATGA	1292
Db	1027	CATAGATGTCATAGATAAGGAACCTAATTCACCTTTTTCAGATCTCTATTTGAAGGATGA	1086
Qy	1293	ATATATTTGCCAGAGCTGGATGACCTCTGAGGGAATATGCTGTCCTCATCTACTAGA	1352
Db	1087	ATATATTTGCCAGAGCTGGATGACCTCTGAGGGAATATGCTGTCCTCATCTACTAGA	1146
Qy	1353	TCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGCAATATATTTATCCAGTAGA	1412
Db	1147	TCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGCAATATATTTATCCAGTAGA	1206
Qy	1413	AGATGATGTTATGGAAGGCGAGACTCAATGAGTCACTGCTGATTCCTGAGCGCACT	1472
Db	1207	AGATGATGTTATGGAAGGCGAGACTCAATGAGTCACTGCTGATTCCTGAGCGCACT	1266
Qy	1473	AATTTATCTAGAGAAACAACAGACATCTGGATTAATATCCATGACATCTTTTATGTTT	1532
Db	1267	AATTTATCTAGAGAAACAACAGACATCTGGATTAATATCCATGACATCTTTTATGTTT	1326
Qy	1533	TCCCAAGTCAAGAGAGAAATTTGAGTTATTTTGGCTCTGAAATGCAAAAACAGGTTT	1592
Db	1327	TCCCAAGTCAAGAGAGAAATTTGAGTTATTTTGGCTCTGAAATGCAAAAACAGGTTT	1386
Qy	1593	CGGTCAATTTATCAAAAATTTACATCTPATTTTAAAGGAAGCAAAATATAAACAGTGG	1652
Db	1387	CGGTCAATTTATCAAAAATTTACATCTPATTTTAAAGGAAGCAAAATATAAACAGTGG	1446
Qy	1653	TGGGCTGCTGCTCCAAAGTATTTCAAGTGTCTTATCAAAAGAGAGATAGCAATACCAAG	1712
Db	1447	TGGGCTGCTGCTCCAAAGTATTTCAAGTGTCTTATCAAAAGAGAGATAGCAATACCAAG	1506
Qy	1713	TGGTGAATGGGAAGTTCTTGGCCGCGCATGATCTTAATATCAAAAGTGTGATGAGTCAAG	1772
Db	1507	TGGTGAATGGGAAGTTCTTGGCCGCGCATGATCTTAATATCAAAAGTGTGATGAGTCAAG	1566
Qy	1773	GCTGGTATATTTTGAAGGACCAAGAGACTCCCTTTTAGAGCATCACCTGTAGTGTAG	1832
Db	1567	GCTGGTATATTTTGAAGGACCAAGAGACTCCCTTTTAGAGCATCACCTGTAGTGTAG	1626
Qy	1833	TTACGTAATCTTGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTG	1892
Db	1627	TTACGTAATCTTGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTG	1686
Qy	1893	CATCAGTCAGCACTGTGACTTCTTTTATAAGTAGTATAGTAAACAGAGAATCCACACTG	1952
Db	1687	CATCAGTCAGCACTGTGACTTCTTTTATAAGTAGTATAGTAAACAGAGAATCCACACTG	1746
Qy	1953	TGTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGAAAT	2012
Db	1747	TGTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGAAAT	1806
Qy	2013	TTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTTCTCTGACTATACCTCCAGAAATTTT	2072
Db	1807	TTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTTCTCTGACTATACCTCCAGAAATTTT	1866
Qy	2073	CTCTTTTGAAGTAGTACTACTGATTTTACATGTATGGATGGATGCTCTACAGCCTCATGATCT	2132
Db	1867	CTCTTTTGAAGTAGTACTACTGATTTTACATGTATGGATGGATGCTCTACAGCCTCATGATCT	1926
Qy	2133	ACAGCCTGGAAGAAATATCTTACTGCTGTCTTATATATGATGCTGCTCAGGTGAGTT	2192
Db	1927	ACAGCCTGGAAGAAATATCTTACTGCTGTCTTATATATGATGCTGCTCAGGTGAGTT	1986
Qy	2193	GGTGAATANTCGGTTTAAAGAGTCAAGTATTTTCCGCTTGAATACCCCTAGCCTCTTAGG	2252
Db	1987	GGTGAATANTCGGTTTAAAGAGTCAAGTATTTTCCGCTTGAATACCCCTAGCCTCTTAGG	2046
Qy	2253	TTATGTGGTTGTAGTAGACAAACAGGGGATCTCTGTACCCAGGGCTTAAATTTGAAGG	2312
Db	2047	TTATGTGGTTGTAGTAGACAAACAGGGGATCTCTGTACCCAGGGCTTAAATTTGAAGG	2106
Qy	2313	CGCCTTTAAATATAAATGGGTCAAAATAGAAATTCAGATCAGGTGGAGGACTCCAATA	2372
Db	2107	CGCCTTTAAATATAAATGGGTCAAAATAGAAATTCAGATCAGGTGGAGGACTCCAATA	2166
Qy	2373	TCTAGCTTCTCGATATGATTTCAATTTGACTTTAGATCGTGTGGGCATCCACGGCTGCTTA	2432
Db	2167	TCTAGCTTCTCGATATGATTTCAATTTGACTTTAGATCGTGTGGGCATCCACGGCTGCTTA	2226
Qy	2433	TGGAGGATACCTCTCCCTGATGGCATTAATGAGAGGTGAGATATCTTTCAGGGTGTCTAT	2492
Db	2227	TGGAGGATACCTCTCCCTGATGGCATTAATGAGAGGTGAGATATCTTTCAGGGTGTCTAT	2286
Qy	2493	TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGAACTTATAT	2552
Db	2287	TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGAACTTATAT	2346
Qy	2553	GGGTCACTCTGACCAAGATGAAACAGGGCTATTACTTTAGGATCTGTGGCCATGCAAGCAGA	2612
Db	2347	GGGTCACTCTGACCAAGATGAAACAGGGCTATTACTTTAGGATCTGTGGCCATGCAAGCAGA	2406
Qy	2613	AAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATCGTTTCTGGATGAGATGT	2672
Db	2407	AAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATCGTTTCTGGATGAGATGT	2466
Qy	2673	CCATTTTGCACATACCGATATATTACTGAGTTTTTATGAGGGCTGGAAGCCATATGA	2732
Db	2467	CCATTTTGCACATACCGATATATTACTGAGTTTTTATGAGGGCTGGAAGCCATATGA	2526
Qy	2733	TTTACAGATCTATCTCAGAGAGAGACAGCATTAAGTTCCTGAAATCGGGAGAACATTA	2792
Db	2527	TTTACAGATCTATCTCAGAGAGAGACAGCATTAAGTTCCTGAAATCGGGAGAACATTA	2586
Qy	2793	TGAACTGCACTTTTGGCATCTACCTCAAGAAAACCTTTGGATACGATATGCTGCTTAAA	2852
Db	2587	TGAACTGCACTTTTGGCATCTACCTCAAGAAAACCTTTGGATACGATATGCTGCTTAAA	2646
Qy	2853	AGTGATAT 2860	
Db	2647	AGTGATAT 2654	

RESULT 11	AR631278	Sequence 2 from patent US 6844180.	2671 bp	DNA	linear	PAT 14-FEB-2005
LOCUS	AR631278					
DEFINITION	AR631278					
ACCESSION	AR631278.1	GI:59770920				
VERSION						
KEYWORDS	Unknown.					
SOURCE	Unclassified.					
ORGANISM	1 (bases 1 to 2671)					
REFERENCE	Oi, S., Akineanya, K.O., Riviere, P.J.M. and Junien, J.-L.					
AUTHORS	Serine protease genes related to DpPiv					
TITLE	Patent: US 6844180-A 2 18-JAN-2005;					
JOURNAL	Ferring BV;					
	NLX;					
FEATURES	Location/Qualifiers					
source	1..2671					
	/organism="unknown"					
	/mol_type="genomic DNA"					
ORIGIN	Query Match 80.0%; Score 2495; DB 2; Length 2671;					
	Best Local Similarity 99.9%; Pred. No. 0;					
	Matches 2645; Conservative 0; Mismatches 3; Indels 0; Gaps 0;					
Qy	213	CATGGCAGCAGCAATGGAAACAGACAGACAGCTGGGTGTGAGATATTGTAAGCTTGAAGCTCGGACTG	272			
Db	7	CATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTGTAAGCTTGAAGCTCGGACTG	66			
Qy	273	TGAGGAGAAATTTGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTTGAGCGGTA	332			
Db	67	TGAGGAGAAATTTGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTTGAGCGGTA	126			
Qy	333	TTCTTGAGTCAGCTTAAAGCTGTTCGGATACCGAATAATCATGCGTACATGAT	392			
Db	127	TTCTTGAGTCAGCTTAAAGCTGTTCGGATACCGAATAATCATGCGTACATGAT	186			
Qy	393	GGCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC	452			
Db	187	GGCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC	246			
Qy	453	AGACAGAACTATTACCTTGCCATGTCGTGTGAGAACAGAGAAATACACTGTTTATTC	512			
Db	247	AGACAGAACTATTACCTTGCCATGTCGTGTGAGAACAGAGAAATACACTGTTTATTC	306			
Qy	513	TGAAATTCCTCAAACTATCAATGAGCAGCAGCTTAAATGCTCTCTTGAAGCCCTCTTTT	572			
Db	307	TGAAATTCCTCAAACTATCAATGAGCAGCAGCTTAAATGCTCTCTTGAAGCCCTCTTTT	366			
Qy	573	GGATCTTTTTCAGGCAACACTGGACTATGGAATGTTATCTCGAGAAAGAACTATTAAAG	632			
Db	367	GGATCTTTTTCAGGCAACACTGGACTATGGAATGTTATCTCGAGAAAGAACTATTAAAG	426			
Qy	633	AGAAAGAAACGCATTGGAACAGTCGGAATGCTTCTTACGATTATCACCAAGGAAGTGG	692			
Db	427	AGAAAGAAACGCATTGGAACAGTCGGAATGCTTCTTACGATTATCACCAAGGAAGTGG	486			
Qy	693	AAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACTGTAATAAGATGGAGGGCCAAAGG	752			
Db	487	AAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACTGTAATAAGATGGAGGGCCAAAGG	546			
Qy	753	ATTATGCGAAACAACTTTAAGCCCAATCTAGTGGAACTAGTGTGTCCTCAACATACGGAT	812			
Db	547	ATTATGCGAAACAACTTTAAGCCCAATCTAGTGGAACTAGTGTGTCCTCAACATACGGAT	606			
Qy	813	CGATCCAAATTTATCCCGCTGATCCAGACTGGAATGCTTTTATACATAGCAACGATAT	872			
Db	607	GGATCCAAATTTATCCCGCTGATCCAGACTGGAATGCTTTTATACATAGCAACGATAT	666			
Qy	873	TTGGATATCTAACTCGTAACACAGAGAAAGAGACTCACTTATGTGCACAATGAGCT	932			

667

TTGGATATCTAACTCGTAACACAGAGAAAGAGACTCACTTATGTGCACAATGAGCT

726

933

AGCCAAACATGGAAGAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCTCAAGAAGA

992

727

AGCCAAACATGGAAGAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCTCAAGAAGA

786

993

ATTTGATAGATATTCTCGCTATTGCTGCTGTCGAAAGCTGAAACAACTCCCAAGTGGTG

1052

787

ATTTGATAGATATTCTCGCTATTGCTGCTGTCGAAAGCTGAAACAACTCCCAAGTGGTG

846

1053

TAAATTTCTTAGAATTTCTATATGAAGAAATGATGATCTGAGGTGGAATTTATTCATGT

1112

847

TAAATTTCTTAGAATTTCTATATGAAGAAATGATGATCTGAGGTGGAATTTATTCATGT

906

1113

TACATCCCTTATGTTGGAAACAAGGAGGAGATCATTTCCGTTATCTTAAACAGGTATC

1172

907

TACATCCCTTATGTTGGAAACAAGGAGGAGATCATTTCCGTTATCTTAAACAGGTATC

966

1173

AGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGAAGGAAGGAT

1232

967

AGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGAAGGAAGGAT

1026

1233

CATAGATCTCATAGATAAGGAACATAATCAACCTTTTGGAGATCTATTTGAAGAGTGA

1292

1027

CATAGATCTCATAGATAAGGAACATAATCAACCTTTTGGAGATCTATTTGAAGAGTGA

1086

1293

ATATATGCCAGAGTGGACTCTGAGGAGGAAATATGCTTGGTCCATCTACTACTAGA

1352

1087

ATATATGCCAGAGTGGAGTGGACTCTGAGGAGGAAATATGCTTGGTCCATCTACTACTAGA

1146

1353

TCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGAAATTAATTTATCCAGTAGA

1412

1147

TCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGAAATTAATTTATCCAGTAGA

1206

1413

AGATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGAGCGACT

1472

1207

AGATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGAGCGACT

1266

1473

AAATATCTATGAAGAAACAACAGACATCTGGATAAATATCATGACATCTTTCAATGTTTT

1532

1267

AAATATCTATGAAGAAACAACAGACATCTGGATAAATATCATGACATCTTTCAATGTTTT

1326

1533

TCCCAAAAGTCACGAAGAGGAAAATTGAGTTTATTTTTCCTCTGAAATGCAAAAACAGTTTT

1592

1327

TCCCAAAAGTCACGAAGAGGAAAATTGAGTTTATTTTTCCTCTGAAATGCAAAAACAGTTTT

1386

1593

CGCTCATTTATACAAAATTAATCTATTTTAAAGGAAGCAAAATATAAAGATCCAGTGG

1652

1387

CGCTCATTTATACAAAATTAATCTATTTTAAAGGAAGCAAAATATAAAGATCCAGTGG

1446

1653

TGGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATTTACCAG

1712

1447

TGGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATTTACCAG

1506

1713

TGTTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATAATCCAGTTGATGAGTCAGAAG

1772

1507

TGTTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATAATCCAGTTGATGAGTCAGAAG

1566

1773

GCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAG

1832

1567

GCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAG

1626

1833

TTACGTAAATCTCGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACTTTCTGCTG

1892

1627

TTACGTAAATCTCGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACTTTCTGCTG

1686

1893

CATCAGTCAGCACTGTGACTCTTTTATAGTAAGTATAGTAACCAAGGAATCCACACTG

1952

1687

CATCAGTCAGCACTGTGACTCTTTTATAGTAAGTATAGTAACCAAGGAATCCACACTG

1746

1953

TGTGTCCCTTTTAAAGCTATCAAGTCTCGGAAGATGACCCCACTTTGCAAAAACAAAGGAAT

2012

1747

TGTGTCCCTTTTAAAGCTATCAAGTCTCGGAAGATGACCCCACTTTGCAAAAACAAAGGAAT

1806

Qy	1053	TAAATTTCTTAGAATTTCTATATGAAGAAAAATGATGAATCTGAGGTGAAAAATATTATCATGT	1112
Db	847	TAATAATCTTAGAATTTCTATATGAAGAAAAATGATGAATCTGAGGTGAAAAATATTATCATGT	906
Qy	1113	TACATCCCTATATGTTGGAAAAAAGAGAGGCGAGATTCAATTCGGTTATCCTAAAAACAGTAC	1172
Db	907	TACATCCCTATGTTGGAAAAAAGAGAGGCGAGATTCAATTCGGTTATCCTAAAAACAGTAC	966
Qy	1173	AGCAATCTTAAGTACATTTTAAAGATGTCAGAAATAATGATGATCTGAAGTGAAGGAAGAT	1232
Db	967	AGCAATCTTAAAGTACATTTTAAAGATGTCAGAAATAATGATGATCTGAAGTGAAGGAAGAT	1026
Qy	1233	CATTAGATGTCATAGATAAGGAACAAATTCACCTTTTGAGATTCCTATTTGAAGAGATTGA	1292
Db	1027	CATTAGATGTCATAGATAAGGAACAAATTCACCTTTTGAGATTCCTATTTGAAGAGATTGA	1086
Qy	1293	ATATATTTGCCAGAGCTGGATGGACTCCTCAGGAGAAAAATATGCTTGGTCCATCTACTAGA	1352
Db	1087	ATATATTTGCCAGAGCTGGATGGACTCCTCAGGAGAAAAATATGCTTGGTCCATCTACTAGA	1146
Qy	1353	TGGTCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCAAGTAGA	1412
Db	1147	TGGTCTCCAGACTCGCCTCAGATAGTGTGATCTCACCTGAATTTATTTATCCCAAGTAGA	1206
Qy	1413	AGATGATGTTATGGAAGGCGAGAGACTCAATGAGTCAGTCGCCTGATTTCTGTGAGCCCACT	1472
Db	1207	AGATGATGTTATGGAAGGCGAGAGACTCAATGAGTCAGTCGCCTGATTTCTGTGAGCCCACT	1266
Qy	1473	AAATATCTATGAAGAAAAACAAGACATCTGGATTAATATTCATGACATCTTTCATGTTTT	1532
Db	1267	AAATATCTATGAAGAAAAACAAGACATCTGGATTAATATTCATGACATCTTTCATGTTTT	1326
Qy	1533	TCCCCAAAGTCAAGAGAGGAATTTAGTTTATTTTTTGGCTCTGATGCAAAACAGGTTTT	1592
Db	1327	TCCCCAAAGTCAAGAGAGGAAATTTAGTTTATTTTTTGGCTCTGATGCAAAACAGGTTTT	1386
Qy	1593	CGGTCAATTTATACAAAATTACATCTATTTTTTAAAGGAAAGCAAAATAFAAACGATCCAGTGG	1652
Db	1387	CGGTCAATTTATACAAAATTACATCTATTTTTTAAAGGAAAGCAAAATAFAAACGATCCAGTGG	1446
Qy	1653	TGGGCTGCTGCTCCAAAGTGATTTCAAAGTGTCTTATCAAAGAGGAGATAGCAATTTACCG	1712
Db	1447	TGGGCTGCTGCTCCAAAGTGATTTCAAAGTGTCTTATCAAAGAGGAGATAGCAATTTACCG	1506
Qy	1713	TGGTGAATGGGAAGTCTTGGCCGCGATGGAATCTAATATCCAAGTTGATGAGTCAAGAG	1772
Db	1507	TGGTGAATGGGAAGTCTTGGCCGCGATGGAATCTAATATCCAAGTTGATGAGTCAAGAG	1566
Qy	1773	GCTGTGATATTTTGAAGCCCAAAAGACTCCCTTTTAGAGCATCACTGTGATCAGTACAG	1832
Db	1567	GCTGTGTATATTTTGAAGGCCCAAAAGACTCCCTTTTAGAGCATCACTGTGATCAGTACAG	1626
Qy	1833	TTACGTAAATCTGGAGAGGTCAAGAGCTGACTGACCGTGGCTACTCAATTTCTTGGCTG	1892
Db	1627	TTACGTAAATCTGGAGAGGTCAAGAGCTGACTGACCGTGGCTACTCAATTTCTTGGCTG	1686
Qy	1893	CATCAGTCAAGCATGTGACTCTTTTAAAGTAAGTATAGTAAACAGAGGAATCCAACATG	1952
Db	1687	CATCAGTCAAGCATGTGACTCTTTTAAAGTAAGTATAGTAAACAGAGGAATCCAACATG	1746
Qy	1953	TGTGTCCCTTTACAAGCTATCAAGTCTGGAAGTACCCAACTTTGCAAAAACAAAGGAATT	2012
Db	1747	TGTGTCCCTTTACAAGCTATCAAGTCTGGAAGTACCCAACTTTGCAAAAACAAAGGAATT	1806
Qy	2013	TTGGGCCACCATTTTGGATTCAGAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTT	2072
Db	1807	TTGGGCCACCATTTTGGATTCAGAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTT	1866
Qy	2073	CTCTTTTGAAGTACTACTGGAATTTATTTGATGGGATGCTCTAACAGCCTCATGATCT	2132
Db	1867	CTCTTTTGAAGTACTACTGGAATTTATTTGATGGGATGCTCTAACAGCCTCATGATCT	1926
Qy	2133	ACAGCCTGGAAAGAAATATCCTATCTGTGCTGTTTCATATATGTGGTCTCTCAGGTGCAGTT	2192

Db	1927	ACAGCCTGGAAAGAAATATCCTACTGTGCTGTTCTATATATGTTGGTCTCAGGTCAGTT	1986
Qy	2193	GGTGAATATCCGTTTAAAGCAGTCAAGTATTTCCGCTTGAATACCTAGACCTCTCTAGG	2252
Db	1987	GGTGAATATCCGTTTAAAGCAGTCAAGTATTTCCGCTTGAATACCTAGACCTCTCTAGG	2046
Qy	2253	TTATGTGGTTGTAGTAGAGACAAGGGAGTCTGTCCAGGAGGCTTAAATTTTGAAGG	2312
Db	2047	TTATGTGGTTGTAGTAGAGACAAGGGAGTCTGTCCAGGAGGCTTAAATTTTGAAGG	2106
Qy	2313	GGCCTTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAAGGCTCCAATA	2372
Db	2107	GGCCTTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAAGGCTCCAATA	2166
Qy	2373	TCTAGCTTCTCGATATGATTTTCATTGAATTAGATCGTGTGGGATCCAAGGCTGGTCTTA	2432
Db	2167	TCTAGCTTCTCGATATGATTTTCATTGAATTAGATCGTGTGGGATCCAAGGCTGGTCTTA	2226
Qy	2433	TGGAGGATACCTCTCCCTCATGGCATTAAATCAGAGGTTCAGATATCTTCAGGGTTGCTAT	2492
Db	2227	TGGAGGATACCTCTCCCTCATGGCATTAAATCAGAGGTTCAGATATCTTCAGGGTTGCTAT	2286
Qy	2493	TGCTGGGGCCCAAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACTGTTATAT	2552
Db	2287	TGCTGGGGCCCAAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACTGTTATAT	2346
Qy	2553	GGGTCAACCTGACAGNATGACAGGGCTATTACTTTAGATCTGTGGCCATGCAAGCAGA	2612
Db	2347	GGGTCAACCTGACAGNATGACAGGGCTATTACTTTAGATCTGTGGCCATGCAAGCAGA	2406
Qy	2613	AAAGTTCCCTCTTGAAACCAATCGTTTACTGCTCTTACATCGTTTCTGTGATGAGATGT	2672
Db	2407	AAAGTTCCCTCTTGAAACCAATCGTTTACTGCTCTTACATCGTTTCTGTGATGAGATGT	2466
Qy	2673	CAATTTTGCAATACAGTATATTACTGTAGTCTTTTGTAGGAGGCTGGAAAGCCATATGA	2732
Db	2467	CAATTTTGCAATACAGTATATTACTGTAGTCTTTTGTAGGAGGCTGGAAAGCCATATGA	2526
Qy	2733	TTTACAGATCTATCCTCAGGAGAGACACAGCATTAAGAGTTCCTGATCCGGAGAACCTTA	2792
Db	2527	TTTACAGATCTATCCTCAGGAGAGACACAGCATTAAGAGTTCCTGATCCGGAGAACCTTA	2586
Qy	2793	TGAATCGCATCTTTTGACTACCTTCAGAGAAACCTTGGATCAAGTATTTGCTGCTCTAAA	2852
Db	2587	TGAATCGCATCTTTTGACTACCTTCAGAGAAACCTTGGATCAAGTATTTGCTGCTCTAAA	2646
Qy	2853	AGTGATAT 2860	
Db	2647	AGTGATAT 2654	

RESULT 13

AB448400

LOCUS

2797 bp

DNA

from patent US 6673549.

linear

PAT 20-FEB-2004

```

RESULT 13
AR448400
LOCUS       AR448400
DEFINITION Sequence 1103 from patent US 6673549.
ACCESSION  AR448400
VERSION     AR448400.1 GI:42676724
KEYWORDS
SOURCE      .
            Unknown.
ORGANISM    Unclassified.
            1 (bases 1 to 2797)
            Furness,J.M. and Buchbinder,J.L.
            Genes expressed in C3A liver cell cultures treated with steroids
            TITLE
            Patent: US 6673549-A 1103 06-JAN-2004;
            INCYTE Corporation; Palo Alto, CA
            FEATURES
            source
            1..2797
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN

```

Query Match									
Best Local Similarity 74.2%; Score 2315; DB 2; Length 2797;									
Matches 2785; Conservative 99.8%; Pred. No. 0; Mismatches 1; Indels 5; Gaps 2;									
Qy	335	CCTGGAGTCAGCTTAAAGCTGCTTCCGATACAGAAAATATCATGGCTACATGATGG	394	Db	1021	AGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAATTTATTTATCCAGT	1080		
Db	1	CCTGGAGTCAGCTTAAAGCTGCTTCCGATACAGAAAATATCATGGCTACATGATGG	60	Qy	1410	AGAAGATGATGTTATGGAAAGCGCAGAGACTCATTTAGTCAGTGCCTGATTTCTGTGACGCC	1469		
Qy	395	CTAAGGCCACCATCATGATTTCAATGTTTGTGAAGAGGAATGATCCAGATGCACTTCATTCAG	454	Db	1081	AGAAGATGATGTTATGGAAAGCGCAGAGACTCATTTAGTCAGTGCCTGATTTCTGTGACGCC	1140		
Db	61	CTAAGGCCACCATCATGATTTCAATGTTTGTGAAGAGGAATGATCCAGATGCACTTCATTCAG	120	Qy	1470	ACTAAATTTATCTATGAAGAAAACAACAGACATCTCGGATAAAATATCATGACATCTTTTCAATGT	1529		
Qy	455	ACAGAAATCTATTACCTTGCCATGCTCGTGAGAACAGAGAAAATACACATGTTTATTCGT	514	Db	1141	ACTAAATTTATCTATGAAGAAAACAACAGACATCTCGGATAAAATATCCATGACATCTTTTCAATGT	1200		
Db	121	ACAGAAATCTATTACCTTGCCATGCTCGTGAGAACAGAGAAAATACACATGTTTATTCGT	180	Qy	1530	TTTTTCCCAAAAGTCCAGAAAGGAAATTTGAGTTTATTTTTTTCCTCTGAAATGCAAAACAGG	1589		
Qy	515	AAATTTCCAAAACACTATCAATAGACGAGCTCTTAATGCTCTCTTGGAAAGCCTCTTTTGG	574	Db	1201	TTTTTCCCAAAAGTCCAGAAAGGAAATTTGAGTTTATTTTTTTCCTCTGAAATGCAAAACAGG	1260		
Db	181	AAATTTCCAAAACACTATCAATAGACGAGCTCTTAATGCTCTCTTGGAAAGCCTCTTTTGG	240	Qy	1590	TTTTCCGTCATTTATACAAAATTAATCATCTATTTTTAAAGGAAAAGCAAAATATAAAACGATCCAG	1649		
Qy	575	ATCTTTTTCAGGCAACACTGGAATATGGAATGATTTCTCGAAGAAAGAACTATTAAAGAG	634	Db	1261	TTTTCCGTCATTTATACAAAATTAATCATCTATTTTTAAAGGAAAAGCAAAATATAAAACGATCCAG	1320		
Db	241	ATCTTTTTCAGGCAACACTGGAATATGGAATGATTTCTCGAAGAAAGAACTATTAAAGAG	300	Qy	1650	TGGTGGGCTGCTGCTCCAAAGTGATTTTCAAGTGTCTTATCAAGAGAGAGATAGCAATTTAC	1709		
Qy	635	AAAGAAAACGATTTGGAACAGTCCGAAATTCCTTTACGATATCAACCAAGGAAGTGGAA	694	Db	1321	TGGTGGGCTGCTGCTCCAAAGTGATTTTCAAGTGTCTTATCAAGAGAGAGATAGCAATTTAC	1380		
Db	301	AAAGAAAACGATTTGGAACAGTCCGAAATTCCTTTACGATATCAACCAAGGAAGTGGAA	360	Qy	1710	CAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTGTGATGAAGTCAG	1769		
Qy	695	CATTTCGTGTTTCAAGCCGGTAGTGGAAATTTATCACTAAAGATGGAGGGCCACAAGGAT	754	Db	1381	CAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTGTGATGAAGTCAG	1440		
Db	361	CATTTCGTGTTTCAAGCCGGTAGTGGAAATTTATCACTAAAGATGGAGGGCCACAAGGAT	420	Qy	1770	AAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTTGAAGCATCACCTGTACGTAGT	1829		
Qy	755	TTACGCAACAACTTTAAGGCCCAATCTAGTGGAACCTAGTTGTGCCCAACATACGGATGG	814	Db	1441	AAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTTGAAGCATCACCTGTACGTAGT	1500		
Db	421	TTACGCAACAACTTTAAGGCCCAATCTAGTGGAACCTAGTTGTGCCCAACATACGGATGG	480	Qy	1830	CAGTTCAGTAAATCCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACTTTCTTG	1889		
Qy	815	ATCCAAATATATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATTT	874	Db	1501	CAGTTCAGTAAATCCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACTTTCTTG	1560		
Db	481	ATCCAAATATATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATTT	540	Qy	1890	CTGCATCAGTCAGCACTGTGACTTTCTTTTATAAGTAAGTATAGTAAACCAAGAAATCCACA	1949		
Qy	875	GGATATCTAACTGCTGTAACAGAGAAAGAGAGACTCACTATGTGCACAATGAGCTAG	934	Db	1561	CTGCATCAGTCAGCACTGTGACTTTCTTTATAAGTAAGTATAGTAAACCAAGAAATCCACA	1620		
Db	541	GGATATCTAACTGCTGTAACAGAGAAAGAGAGACTCACTATGTGCACAATGAGCTAG	600	Qy	1950	CTGTGTGCTCCCTTTACAAGCTATCAAGTCTCTGAAATGACCAACTTGCAAAACAAGGA	2009		
Qy	935	CCAACTGGGAAGAGATGCCAGATCAGCTGGAGTGCCTACCTTTGTTTCTCCAGAAAGAT	994	Db	1621	CTGTGTGCTCCCTTTACAAGCTATCAAGTCTCTGAAATGACCAACTTGCAAAACAAGGA	1680		
Db	601	CCAACTGGGAAGAGATGCCAGATCAGCTGGAGTGCCTACCTTTGTTTCTCCAGAAAGAT	660	Qy	2010	ATTTTGGGGCCACATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAAT	2069		
Qy	995	TTGATAGATATCTGGCTATTTGGTGTGTCMAAAGCTGAAACAACTCCAGTGGTGGTA	1054	Db	1681	ATTTTGGGGCCACATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAAT	1740		
Db	661	TTGATAGATATCTGGCTATTTGGTGTGTCMAAAGCTGAAACAACTCCAGTGGTGGTA	720	Qy	2070	TTTTCTCTTTTGAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAGGCTCATGA	2129		
Qy	1055	AAATTTCTTGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAATTTATTCATGTTA	1114	Db	1741	TTTTCTCTTTTGAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAGGCTCATGA	1800		
Db	721	AAATTTCTTGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAATTTATTCATGTTA	780	Qy	2130	TCTACAGCTCGGAAAAGAAATATCCTACTGTGCTGTTTCAATATATGGTGGTCTCTCAGGTGCA	2189		
Qy	1115	CATCCCTATGTTGGAAAACAGG-AGGCGAGATTCAATTCGGTTATCTTAAACAGGTACA	1173	Db	1801	TCTACAGCTCGGAAAAGAAATATCCTACTGTGCTGTTTCAATATATGGTGGTCTCTCAGGTGCA	1860		
Db	781	CATCCCTATGTTGGAAAACAGGAGGCGCAGATTCATTCGGTTATCTTAAACAGGTACA	840	Qy	2190	GTTTGGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTTCCGCTTTGAATACCTTAGGCTCTCT	2249		
Qy	1174	GCAATTCCTAAAGTCACTTTTAAAGATGTCAGAAATTAATGATTTGATGCTGAAGGAAGATC	1233	Db	1861	GTTTGGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTTCCGCTTTGAATAACCTTAGGCTCTCT	1920		
Db	841	GCAATTCCTAAAGTCACTTTTAAAGATGTCAGAAATTAATGATTTGATGCTGAAGGAAGATC	900	Qy	2250	AGGTTATGTGGTGTGTAGTGATAGACAAACAGGGGATCTGTGTCAACCGAGGGCTTAAATTTGA	2309		
Qy	1234	ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTTCTATTTGAAGGAGTTGAA	1293	Db	1921	AGGTTATGTGGTGTGTAGTGATAGACAAACAGGGGATCTGTGTCAACCGAGGGCTTAAATTTGA	1980		
Db	901	ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTTCTATTTGAAGGAGTTGAA	960	Qy	2310	AGGCGCTTTTAAATATAAAATGGGTCAAATAGAAATTTGACGATCAGTGGGAAGACTCCA	2369		
Qy	1294	TATATTTGCCAGAGCTGGATGGACCTCTGAGGGAATA-----TGCTTGGTCCATCTACT	1349	Db	1981	AGGCGCTTTTAAATATAAAATGGGTCAAATAGAAATTTGACGATCAGTGGGAAGACTCCA	2040		
Db	961	TATATTTGCCAGAGCTGGATGGACCTCTGAGGGAATA-----TGCTTGGTCCATCTACT	1020	Qy	2370	ATATCTAGCTTCTCGATATGATTTTCAATGACTTATAGATCGTGTGGGCAATCCACGGCTGGTC	2429		
Qy	1350	AGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCACTGGAATTTATTTATCCCAAGT	1409	Db	2041	ATATCTAGCTTCTCGATATGATTTTCAATGACTTATAGATCGTGTGGGCAATCCACGGCTGGTC	2100		
Db				Qy	2430	CTATGAGAGATACCTCTCCCTGATGGCAATTAATCCAGGTCAGATATCTTCTCAGGGTTCG	2489		

Db	2101	CTATGGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTTCAGATATCTTCAGGGTTGC	2160
Qy	2490	TATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAACTTA	2549
Db	2161	TATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAACTTA	2220
Qy	2550	TATGGGTCACTTCAGCAGAGATGAACAGGGCTATTAATCTAGATCTGTGGCCATCAAGC	2609
Db	2221	TATGGGTCACTTCAGCAGAGATGAACAGGGCTATTAATCTAGATCTGTGGCCATCAAGC	2280
Qy	2610	AGAAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTGATGAGAA	2669
Db	2281	AGAAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTGATGAGAA	2340
Qy	2670	TGTCATTTTGGCATACATACAGTATATTAATCTAGATCTTGTAGGGCTTGAAGCCATA	2729
Db	2341	TGTCATTTTGGCATACATACAGTATATTAATCTAGATCTTGTAGGGCTTGAAGCCATA	2400
Qy	2730	TGATTTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTTCTGAAATCGGAGAAC	2789
Db	2401	TGATTTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTTCTGAAATCGGAGAAC	2460
Qy	2790	TTATGAATGATCTTTTGGCACTACCTTCAAGAAACCTTGGATCACTGATTTGCTGCTCT	2849
Db	2461	TTATGAATGATCTTTTGGCACTACCTTCAAGAAACCTTGGATCACTGATTTGCTGCTCT	2520
Qy	2850	AAAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCA	2909
Db	2521	AAAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCA	2580
Qy	2910	ATGAGGAGTTTAAATCAACGAAACACAGAAATGATATCATATTTGATACCTGCCAT	2969
Db	2581	ATGAGGAGTTTAAATCAACGAAACACAGAAATGATATCATATTTGATACCTGCCAT	2640
Qy	2970	GTAACATCTACTCTGAAATTAATGTTGGTGCATCGAGGGTCTACGGTTTGTGTAGT	3029
Db	2641	GTAACATCTACTCTGAAATTAATGTTGGTGCATCGAGGGTCTACGGTTTGTGTAGT	2700
Qy	3030	AATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATATATTTCTGAGAGACCA	3089
Db	2701	AATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATATATTTCTGAGAGACCA	2760
Qy	3090	GCAATACCATGAAGATTAATTAATAAAAAA 3120	
Db	2761	GCAATACCATGAAGATTAATTAATAAAAAA 2791	

RESULT 14
AR631281
LOCUS AR631281 4523 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 8 from patent US 6844180.
ACCESSION AR631281
VERSION AR631281.1 GI:59770924
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4523)
AUTHORS O.I.S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 8 18-JAN-2005;
Ferring BV;
NLX;
FEATURES
source Location/Qualifiers
1..4523
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 68.2%; Score 2128; DB 2; Length 4523;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	AAGTGTAAAGCCTCCGAGGCCAAGGCGCTGTCTACTGCGCGCGCTGCTTCTTAGTGCCG	60
Db	1	AAGTGTAAAGCCTCCGAGGCCAAGGCGCTGTCTACTGCGCGCGCTGCTTCTTAGTGCCG	60
Qy	61	CGTTCCGCGCTGGGTTGTCAACCGCGCGCGCGCGCGAGGAGGACATCTGCAACACGAGACG	120
Db	61	CGTTCCGCGCTGGGTTGTCAACCGCGCGCGCGCGCGAGGAGGACATCTGCAACACGAGACG	120
Qy	121	GAGTGGAGCGCGCAGCATGAAGCGCGCGCGCGCGCTCATAGCCACGTCGGGACGG	180
Db	121	GAGTGGAGCGCGCAGCATGAAGCGCGCGCGCGCGCTCATAGCCACGTCGGGACGG	180
Qy	181	TCCGGCGCGCGCGCGGAGGAAATGCAACATGCGCAGCAATGGAACACAGAACAG	240
Db	181	TCCGGCGCGCGCGCGGAGGAAATGCAACATGCGCAGCAATGGAACACAGAACAG	240
Qy	241	CTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGATATTTGAATCACAGGATCGG	300
Db	241	CTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGATATTTGAATCACAGGATCGG	300
Qy	301	CCTAAATTTGGAGCCTTTTATGTTGAGCGGTATTTCTTGGAGTCAGCTTAAAAAGCTGCTT	360
Db	301	CCTAAATTTGGAGCCTTTTATGTTGAGCGGTATTTCTTGGAGTCAGCTTAAAAAGCTGCTT	360
Qy	361	GCCGATACAGAAATATCATGCTACATGATGCTTAAGGACCAACATGATTTTCATGCTTT	420
Db	361	GCCGATACAGAAATATCATGCTACATGATGCTTAAGGACCAACATGATTTTCATGCTTT	420
Qy	421	GTGAGAGGAATGATTCAGATGGACCTCATTCAGACAGAACTCTATTACCTTGGCATGCT	480
Db	421	GTGAGAGGAATGATTCAGATGGACCTCATTCAGACAGAACTCTATTACCTTGGCATGCT	480
Qy	481	GSTGAGAACAGAGAAATATCACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGACGA	540
Db	481	GSTGAGAACAGAGAAATATCACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGACGA	540
Qy	541	GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Db	541	GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Qy	601	GGAATGTAATTCGAGAGAGAACTATTAAGAGAGAGAGAAACGCAATTTGGAACAGTCGGA	660
Db	601	GGAATGTAATTCGAGAGAGAGAACTATTAAGAGAGAGAGAAACGCAATTTGGAACAGTCGGA	660
Qy	661	ATTGCTTCTTACGATTTATCAACAGAGAGTGGACATTTCTGTTTCAAGCGGTAGTGA	720
Db	661	ATTGCTTCTTACGATTTATCAACAGAGAGTGGACATTTCTGTTTCAAGCGGTAGTGA	720
Qy	721	ATTTATCAGTAAAGATGGAGGCGCACAGGATTTACGCAACCACTTTAAAGGCCCAAT	780
Db	721	ATTTATCAGTAAAGATGGAGGCGCACAGGATTTACGCAACCACTTTAAAGGCCCAAT	780
Qy	781	CTAGTGGAAACTAGTTGTCCCAACATACGATGGATCCAAATTTATGCCCGCTGATCCA	840
Db	781	CTAGTGGAAACTAGTTGTCCCAACATACGATGGATCCAAATTTATGCCCGCTGATCCA	840
Qy	841	GACTGGAATGCTTTTATACATAGCAACGATATTTGGATATCTAACTCGTAACAGAGAA	900
Db	841	GACTGGAATGCTTTTATACATAGCAACGATATTTGGATATCTAACTCGTAACAGAGAA	900
Qy	901	GAAAGGAGACTCACTTATGTGCACAATAGCTAGCCAACTGAGAGAGAGATGCCAGATCA	960
Db	901	GAAAGGAGACTCACTTATGTGCACAATAGCTAGCCAACTGAGAGAGAGATGCCAGATCA	960
Qy	961	GCTGGAGTGGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
Db	961	GCTGGAGTGGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
Qy	1021	TGTCCAAAGCTGAAACAACTCCAGTGTGTGTAATAATTTCTTAGAATTTCTATATGAAGAA	1080
Db	1021	TGTCCAAAGCTGAAACAACTCCAGTGTGTGTAATAATTTCTTAGAATTTCTATATGAAGAA	1080
Qy	1081	AATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCTATGTTGGAAACAGGAGG	1140

Db	1081	AATGATGAATCTGAGGTGGAATATTATTCATGTATACATCCCTATGTTGGAAACAGAGG	1140
Qy	1141	GCAGATTCAATCCGTTATATCCATAAACAGGTACACGAATCTTAAGTCACCTTTAAGATG	1200
Db	1141	GCAGATTCAATCCGTTATATCCATAAACAGGTACACGAATCTTAAGTCACCTTTAAGATG	1200
Qy	1201	TCAGAAATATGATTGATCTGAAGGAGGATCATAGATGTCATAGATGAAGAACTAAAT	1260
Db	1201	TCAGAAATATGATTGATCTGAAGGAGGATCATAGATGTCATAGATGAAGAACTAAAT	1260
Qy	1261	CAACCTTTTGAGATCTATTTGAAGGAGTTGAATATATTCGCCAGAGCTGGATGCTCT	1320
Db	1261	CAACCTTTTGAGATCTATTTGAAGGAGTTGAATATATTCGCCAGAGCTGGATGCTCT	1320
Qy	1321	GAGGAAATATGCTTGCTCATCTAGATCGCTCCAGACTCGCCTACAGATAGTG	1380
Db	1321	GAGGAAATATGCTTGCTCATCTAGATCGCTCCAGACTCGCCTACAGATAGTG	1380
Qy	1381	TTGATCTCACCTGAATATTATATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC	1440
Db	1381	TTGATCTCACCTGAATATTATATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC	1440
Qy	1441	ATTGAGTCAGTCCCTGATTCGTGACGCCACTAATTTATCTATGAAGAAACACAGACATC	1500
Db	1441	ATTGAGTCAGTCCCTGATTCGTGACGCCACTAATTTATCTATGAAGAAACACAGACATC	1500
Qy	1501	TGGATAATATCCATGATCATCTTTTCATGTTTTTCCCAAAAGTCACGAAGAGGAATTTGAG	1560
Db	1501	TGGATAATATCCATGATCATCTTTTCATGTTTTTCCCAAAAGTCACGAAGAGGAATTTGAG	1560
Qy	1561	TTTATTTTGTCTCTGAATCAAAACAGGTTTCGGTCATTTATATACAAATTTACATCTATT	1620
Db	1561	TTTATTTTGTCTCTGAATCAAAACAGGTTTCGGTCATTTATATACAAATTTACATCTATT	1620
Qy	1621	TTAAAGGAAAGCAATATATAACGATCCAGTGTGGGCTGCTCCAGTGATTTCAAG	1680
Db	1621	TTAAAGGAAAGCAATATATAACGATCCAGTGTGGGCTGCTCCAGTGATTTCAAG	1680
Qy	1681	TGTCCTATCAAGAGAGATAGCAATTAACAGTGTGATGGAGTTCTTGGCCGGCAT	1740
Db	1681	TGTCCTATCAAGAGAGATAGCAATTAACAGTGTGATGGAGTTCTTGGCCGGCAT	1740
Qy	1741	GGATCTAATATCCAAGTTCAGTTCAGAGGCTGGTATATTTTGAAGGCACCAAGAC	1800
Db	1741	GGATCTAATATCCAAGTTCAGTTCAGAGGCTGGTATATTTTGAAGGCACCAAGAC	1800
Qy	1801	TCCCTTTAGAGGATCACTGTACGTAGTCAGTTACGTAATCTCGAGAGGTGACAAGG	1860
Db	1801	TCCCTTTAGAGGATCACTGTACGTAGTCAGTTACGTAATCTCGAGAGGTGACAAGG	1860
Qy	1861	CTGACTGACGGTGCTACTCAATTTGCTGCTGCACTGAGCTGAGCTCTCTTTATA	1920
Db	1861	CTGACTGACGGTGCTACTCAATTTGCTGCTGCACTGAGCTGAGCTCTCTTTATA	1920
Qy	1921	AGTAAGTATAGTAAACAAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCT	1980
Db	1921	AGTAAGTATAGTAAACAAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCT	1980
Qy	1981	GAAGATGACCCAACTTGCAAAACAAAGGAAATTTTGGGCCACCAATTTTGAATTCAGAGCT	2040
Db	1981	GAAGATGACCCAACTTGCAAAACAAAGGAAATTTTGGGCCACCAATTTTGAATTCAGAGCT	2040
Qy	2041	CCTCTCTGACTATACCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTTACA	2100
Db	2041	CCTCTCTGACTATACCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTTACA	2100
Qy	2101	TTGATGGAGTCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATCTCTACTG	2160
Db	2101	TTGATGGAGTCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATCTCTACTG	2160
Qy	2161	CTGTTTCATATATGGTGTGTC	2179
Db	2161	CTGTTTCATATATGGTGTGTC	2179
RESULT 15			
AX608731			
LOCUS			
DEFINITION	Sequence 8 from Patent WO0231134.	4523 bp	DNA linear PAT 17-FEB-2003
ACCESSION	AX608731		
VERSION	AX608731.1	GI:28404301	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.		
TITLE	Novel serine protease genes related to dppiv		
JOURNAL	Patent: WO 0231134-A 8 18-APR-2002;		
	Ferring BV (NL)		
FEATURES	Location/Qualifiers		
source	1..4523		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	68.2%;	Score 2128;	DB 2; Length 4523;
Best Local Similarity	99.9%;	Pred. No. 0;	
Matches 2178;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	1	AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCTGCTTCTTAGTGCGG	60
Db	1	AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCTGCTTCTTAGTGCGG	60
Qy	61	CGTTCCGCGCCTGGGTTGTCAACCGCGCCGCCGCGGAGGAGCACTGCAACACAGACCG	120
Db	61	CGTTCCGCGCCTGGGTTGTCAACCGCGCCGCCGCGGAGGAGCACTGCAACACAGACCG	120
Qy	121	GAGTGGAGGGGGCGGCGACATGAAGCGCGCGCGCGCTCCATAGCGCACTGCGGACCG	180
Db	121	GAGTGGAGGGGGCGGCGACATGAAGCGCGCGCGCGCTCCATAGCGCACTGCGGACCG	180
Qy	181	TCCGGCGCGCGCGCGGGGGAAGAAATGCAACATGCGCGCGCGCAATGGAAACAGAACAG	240
Db	181	TCCGGCGCGCGCGCGGGGGAAGAAATGCAACATGCGCGCGCGCAATGGAAACAGAACAG	240
Qy	241	CTGGTGCTTCAGATATTTGAAACTCGGACTCTGAGGAGATATTGAATTCACAGATCGG	300
Db	241	CTGGTGCTTCAGATATTTGAAACTCGGACTCTGAGGAGATATTGAATTCACAGATCGG	300
Qy	301	CCTAAATTGGAGCCTTTTATGTTGAGCGGTTATTCCTGGAGTCACTTAAAGCTGCTT	360
Db	301	CCTAAATTGGAGCCTTTTATGTTGAGCGGTTATTCCTGGAGTCACTTAAAGCTGCTT	360
Qy	361	GCGGATACCGAATAATATCATGGCTACATGATGGCTAAGGCCACCACTGATTTTCTGTTT	420
Db	361	GCGGATACCGAATAATATCATGGCTACATGATGGCTAAGGCCACCACTGATTTTCTGTTT	420
Qy	421	GTGAGAGGAATGATCCAGATGGACCTCATTTGAGCAGAAATCTATTACCTTGGCATGCT	480
Db	421	GTGAGAGGAATGATCCAGATGGACCTCATTTGAGCAGAAATCTATTACCTTGGCATGCT	480
Qy	481	GGTGAGAAACAGAGAAATACACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGAGCA	540
Db	481	GGTGAGAAACAGAGAAATACACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGAGCA	540
Qy	541	GCAGTCTTAATGCTCTCTTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Db	541	GCAGTCTTAATGCTCTCTTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Qy	601	GGAAATGATTTCTCGAGAAAGAACTATTAAGAGAAAGAAACCGCATTTGGAACAGTCGGA	660
Db	601	GGAAATGATTTCTCGAGAAAGAACTATTAAGAGAAAGAAACCGCATTTGGAACAGTCGGA	660

601 GGAATGATATTCTCGAAGAGAACTAATTAAGAGAAAGAAAACGCAATTTGGAACAGTCCGA 660
661 ATTGCTCTTACGATTATCACCAAGGAAGTGAACAATTTCTGTCTTCAAGCCGGTAGTGA 720
661 ATTGCTCTTACGATTATCACCAAGGAAGTGAACAATTTCTGTCTTCAAGCCGGTAGTGA 720
721 ATTATACAGTAAAGATGGAGGCCACAAGGATTTACGCAACAACTTTAAGGCCCAAT 780
721 ATTATACAGTAAAGATGGAGGCCACAAGGATTTACGCAACAACTTTAAGGCCCAAT 780
781 CTAGTGGAACTAGTTGTCCTCCCAACATACGGAATGATCCAAATTTATGCCCGCTGATCCA 840
781 CTAGTGGAACTAGTTGTCCTCCCAACATACGGAATGATCCAAATTTATGCCCGCTGATCCA 840
841 GACTGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACGAGAA 900
841 GACTGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACGAGAA 900
901 GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAAATGGAAGAAAGATGCCAGATCA 960
901 GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAAATGGAAGAAAGATGCCAGATCA 960
961 GCTGGAGTGGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTTGGTG 1020
961 GCTGGAGTGGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTTGGTG 1020
1021 TGTCCTAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTAGAAATTTCTATATGAAGAA 1080
1021 TGTCCTAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTAGAAATTTCTATATGAAGAA 1080
1081 AATGATGAATCTGAGGTGGAAATTTATCATGTTACATCCCTATGTTGGAAACAAAGAGG 1140
1081 AATGATGAATCTGAGGTGGAAATTTATCATGTTACATCCCTATGTTGGAAACAAAGAGG 1140
1141 GCAGATTCATTCGGTTATCTCTAAACAGGTACAGCAATCTCTAAAGTCACTTTTAAGATG 1200
1141 GCAGATTCATTCGGTTATCTCTAAACAGGTACAGCAATCTCTAAAGTCACTTTTAAGATG 1200
1201 TCAGAAATATGATGATGCTGAGGAGGATCATAGATGTCATAGATAAGAACTAAT 1260
1201 TCAGAAATATGATGATGCTGAGGAGGATCATAGATGTCATAGATAAGAACTAAT 1260
1261 CAACCTTTTGAGATTTCTATTTGAAGGATTTGAATATTTGCCAGAGCTGGATGACTCT 1320
1261 CAACCTTTTGAGATTTCTATTTGAAGGATTTGAATATTTGCCAGAGCTGGATGACTCT 1320
1321 GAGGGAATATGCTTGGTCCATCTCTAGATGCTGCCAGCTGCTACAGATGATG 1380
1321 GAGGGAATATGCTTGGTCCATCTCTAGATGCTGCCAGCTGCTACAGATGATG 1380
1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAGATGATGTTATGGAAGGAGAGACTC 1440
1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAGATGATGTTATGGAAGGAGAGACTC 1440
1441 ATTGAGTCAGTGCCTGATTTCTGAGCCCACTAATTTATCTATGAAGAAACACACATC 1500
1441 ATTGAGTCAGTGCCTGATTTCTGAGCCCACTAATTTATCTATGAAGAAACACACATC 1500
1501 TGGATAAATATCCATGACATCTTTCAATGTTTTTCCCAAGTCAAGAGAGAAATTTGAG 1560
1501 TGGATAAATATCCATGACATCTTTCAATGTTTTTCCCAAGTCAAGAGAGAAATTTGAG 1560
1561 TTTATTTTGGCTCTGAAATGCAAAAGGTTTCCGTCATTTATACAAATATCACTAAT 1620
1561 TTTATTTTGGCTCTGAAATGCAAAAGGTTTCCGTCATTTATACAAATATCACTAAT 1620
1621 TTAAGGAAGCAATATAAAGATCCAGTGGTGGCTGCTCTCCAAAGTATTTCAAG 1680
1621 TTAAGGAAGCAATATAAAGATCCAGTGGTGGCTGCTCTCCAAAGTATTTCAAG 1680
1681 TGTCTCTATCAAGAGGAGATAGCAATTTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
1681 TGTCTCTATCAAGAGGAGATAGCAATTTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740

1741 GGATCTAATATCCAGTTGATGAAGCTCAGAAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
1741 GGATCTAATATCCAGTTGATGAAGCTCAGAAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
1801 TCCCTTTTAGAGCATCACCTGTACGTAGTACGTAAATCTCGAAGAGGTGACAAGG 1860
1801 TCCCTTTTAGAGCATCACCTGTACGTAGTACGTAAATCTCGAAGAGGTGACAAGG 1860
1861 CTGACTGACCGTGGCTACTCATCTTCTGCTGATCAGTCAAGCTGCTGATCTTTATA 1920
1861 CTGACTGACCGTGGCTACTCATCTTCTGCTGATCAGTCAAGCTGCTGATCTTTATA 1920
1921 AGTAAGTATAGTAACAGAGAAATCCACACTGCTGTGCTCTTACCAAGCTATCAAGTCT 1980
1921 AGTAAGTATAGTAACAGAGAAATCCACACTGCTGTGCTCTTACCAAGCTATCAAGTCT 1980
1981 GAAGATGACCAACTTTGCAAAACAAAGAAATTTTGGGCCACCATTTGGATTACAGAGT 2040
1981 GAAGATGACCAACTTTGCAAAACAAAGAAATTTTGGGCCACCATTTGGATTACAGAGT 2040
2041 CCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTACA 2100
2041 CCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTACA 2100
2101 TTGATGGGATGCTTACAGGCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTG 2160
2101 TTGATGGGATGCTTACAGGCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTG 2160
2161 CTGTTCTATATGTTGCTG 2179
2161 CTGTTCTATATGTTGCTG 2179

RESULT 16
AR631287 4676 bp DNA linear PAT 14-FEB-2005
LOCUS Sequence 20 from patent US 6844180.
DEFINITION AR631287
ACCESSION AR631287
VERSION AR631287.1 GI:59770934
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 4676)
AUTHORS OI.S., Akinsanya,K.O., Riviere,P.J.M. and Junien,J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 20 18-JAN-2005;
Ferring BV;;
NLX;

FEATURES
source
1..4676
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 68.2%; Score 2128; DB 2; Length 4676;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAGTGTAAAGCTCCGAGGCCAGGCGCTGCTACTGCGCGCTGCTCTTTAGTGGCG 60
DB 1 AAGTGTAAAGCTCCGAGGCCAGGCGCTGCTACTGCGCGCTGCTCTTTAGTGGCG 60
QY 61 CGTTGCGCGCTGGGTGTGTCAACCGCGCGCGCGCGAGGAGCCACTGCAACACGAGCCG 120
DB 61 CGTTGCGCGCTGGGTGTGTCAACCGCGCGCGCGAGGAGCCACTGCAACACGAGCCG 120
QY 121 GAGTGGAGCGCGCAGCATGAGCGCGCGCGCGCGCTCCATAGCCAGTGGGACCG 180
DB 121 GAGTGGAGCGCGCAGCATGAGCGCGCGCGCGCTCCATAGCCAGTGGGACCG 180
QY 181 TCCGGGCGGGCGGGGGAGGAAATGCAACATGGCAGCAGCAATGGAAACAGAAACAG 240

Db 181 TCCGGCGGGCGGGGGAAGGAAAATGCAACATGCGACGACCAATGNNACAGAC 240
Qy 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTCGAATCAAGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTCGAATCAAGATCGG 300
Qy 301 CCTAAATTTGGAGCCCTTTTATGTTGAGCGGTNTCTGAGTCAGCTTAAAAAGCTGCTT 360
Db 301 CCTAAATTTGGAGCCCTTTTATGTTGAGCGGTNTCTGAGTCAGCTTAAAAAGCTGCTT 360
Qy 361 GCCGATACAGAAATATCATGATGCTTAAGGCCACCAATGATTTCAATGTTT 420
Db 361 GCCGATACAGAAATATCATGCTATCATGCTTAAGGCCACCAATGATTTCAATGTTT 420
Qy 421 GTGAGAGGATGATCCGATGGAACCTCATTCAGACAGAAATCTATPACCTTCCCATGCT 480
Db 421 GTGAGAGGATGATCCGATGGAACCTCATTCAGACAGAAATCTATPACCTTCCCATGCT 480
Qy 481 GGTGAGAAACAGAGAAATACATGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540
Db 481 GGTGAGAAACAGAGAAATACATGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540
Qy 541 GCAGTCTTAATGCTCTCTTGGAAAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTGGAAAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Qy 601 GGAATGTAATTTCTGAGAGAGAGAACTATTAAGAGAGAGAGAAACGCAATTTGGAACAGTCGGA 660
Db 601 GGAATGTAATTTCTGAGAGAGAGAACTATTAAGAGAGAGAGAAACGCAATTTGGAACAGTCGGA 660
Qy 661 ATTGCTTCTTACGATATCACTCAAGGAAGTGGAAATTTCTGTTTCAAGCGGTAGTGGGA 720
Db 661 ATTGCTTCTTACGATATCACTCAAGGAAGTGGAAATTTCTGTTTCAAGCGGTAGTGGGA 720
Qy 721 ATTTATCACTGTAAGAGATGAGGCGCCACAGGATTTAGCACAACCTTTAAGGCCCAAT 780
Db 721 ATTTATCACTGTAAGAGATGAGGCGCCACAGGATTTAGCACAACCTTTAAGGCCCAAT 780
Qy 781 CTAGTGAAACTAGTTGTCCCAACATACGGATGGATCCAAATTTAGCCGCTGATCCA 840
Db 781 CTAGTGAAACTAGTTGTCCCAACATACGGATGGATCCAAATTTAGCCGCTGATCCA 840
Qy 841 GACTGGATTCCTTTATACATACGACGATATTTGGATATCTTAACATCGTAAACAGAGAA 900
Db 841 GACTGGATTCCTTTATACATACGACGATATTTGGATATCTTAACATCGTAAACAGAGAA 900
Qy 901 GAAAGGAGACTCACTTATGTCACATGAGCTAGCCAAATGGAAGAGATGCGCAGATCA 960
Db 901 GAAAGGAGACTCACTTATGTCACATGAGCTAGCCAAATGGAAGAGATGCGCAGATCA 960
Qy 961 GCTGGAGTCGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTCGGCTATTTGGTGG 1020
Db 961 GCTGGAGTCGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTCGGCTATTTGGTGG 1020
Qy 1021 TGTCCAAAAGCTGAACAACTCCCAAGTGGTGAATTTCTAGAAATTTCTATATGAAGAA 1080
Db 1021 TGTCCAAAAGCTGAACAACTCCCAAGTGGTGAATTTCTAGAAATTTCTATATGAAGAA 1080
Qy 1081 AATGATGAATCTGAGGTGAAATTTATCATGTTATCATCCCTATGTTGGAAACAGGAGG 1140
Db 1081 AATGATGAATCTGAGGTGAAATTTATCATGTTATCATCCCTATGTTGGAAACAGGAGG 1140
Qy 1141 GCAGATTCATTCGGTTATCCTAAAAACAGGTACAGCAATCTTAAAGTCACTTTTAAGATG 1200
Db 1141 GCAGATTCATTCGGTTATCCTAAAAACAGGTACAGCAATCTTAAAGTCACTTTTAAGATG 1200
Qy 1201 TCAGAAATTAATGATGATGCTCAAGCAAGGATCATAGATGTCATAGATAGGAACTAAT 1260
Db 1201 TCAGAAATTAATGATGATGCTCAAGCAAGGATCATAGATGTCATAGATAGGAACTAAT 1260
Qy 1261 CAACCTTTTGAATTCATTTTGAAGGAGTTCGAATATATTCAGAGCTGGATGCTCT 1320
Db 1261 CAACCTTTTGAATTCATTTTGAAGGAGTTCGAATATATTCAGAGCTGGATGCTCT 1320

Db 1261 CAACCTTTTGAATTCATTTTGAAGGAGTTCGAATATATTCAGAGCTGGATGCTCT 1320
Qy 1321 GAGGAAAAATATGCTTTGGTCCATCTCTAGATCGCTCCAGACTCGCTACAGATAGTG 1380
Db 1321 GAGGAAAAATATGCTTTGGTCCATCTCTAGATCGCTCCAGACTCGCTACAGATAGTG 1380
Qy 1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATTTATGGAAGGAGAGACTC 1440
Db 1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATTTATGGAAGGAGAGACTC 1440
Qy 1441 ATTGAGTCAGTCCCTGATTTCTGTGACGCCACTTAATTTATGATGAAGAAACAAAGACATC 1500
Db 1441 ATTGAGTCAGTCCCTGATTTCTGTGACGCCACTTAATTTATGATGAAGAAACAAAGACATC 1500
Qy 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAAAGTCCAGAGAGGAAATGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAAAGTCCAGAGAGGAAATGAG 1560
Qy 1561 TTTATTTTGGCTCTGATGCAAAACAGGTTTCGTCATTTATACAAATTAATCATTT 1620
Db 1561 TTTATTTTGGCTCTGATGCAAAACAGGTTTCGTCATTTATACAAATTAATCATTT 1620
Qy 1621 TTAAGAGAAACCAATATTAACGATCCAGTGGTGGCTGCTGCTCCAGTGAATTTCAAG 1680
Db 1621 TTAAGAGAAACCAATATTAACGATCCAGTGGTGGCTGCTGCTCCAGTGAATTTCAAG 1680
Qy 1681 TGTCTCTATCAAGAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Db 1681 TGTCTCTATCAAGAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Qy 1741 GGAATCTAATTCGAAGTGAATGAAGTCAAGGCTGGTATATTTTGAAGGCCAACCAAGAC 1800
Db 1741 GGAATCTAATTCGAAGTGAATGAAGTCAAGGCTGGTATATTTTGAAGGCCAACCAAGAC 1800
Qy 1801 TCCCTTTTGAAGCATCACTGTAGTAGTCAGTTACGTAATCTCGAGAGGTGACAAGG 1860
Db 1801 TCCCTTTTGAAGCATCACTGTAGTAGTCAGTTACGTAATCTCGAGAGGTGACAAGG 1860
Qy 1861 CTGATGACCGTGGCTACTCACATTTCTGTCATCAGTCAGCACTGTGACTTCTTTATA 1920
Db 1861 CTGATGACCGTGGCTACTCACATTTCTGTCATCAGTCAGCACTGTGACTTCTTTATA 1920
Qy 1921 AGTAAGTATAGTAACCAAGAGAAATCCACATGTCCTCTTTTACAAGCTATCAAGTCT 1980
Db 1921 AGTAAGTATAGTAACCAAGAGAAATCCACATGTCCTCTTTTACAAGCTATCAAGTCT 1980
Qy 1981 GAAGATGACCCAACTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGT 2040
Db 1981 GAAGATGACCCAACTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGT 2040
Qy 2041 CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTTACA 2100
Db 2041 CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTTACA 2100
Qy 2101 TTGATGCGGATGCTTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTG 2160
Db 2101 TTGATGCGGATGCTTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTG 2160
Qy 2161 CTGTTCTATATGCTGCTC 2179
Db 2161 CTGTTCTATATGCTGCTC 2179

RESULT 17

AX608743

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AX608743 Sequence 20 from Patent WO0231134.
AX608743 GI:28404307

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

linear PAT 17-FEB-2003

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.

AUTHORS Novel serine protease genes related to dplv

TITLE Patent: WO 0231134-A 20 18-APR-2002;

JOURNAL Ferring BV (NL)

FEATURES Location/Qualifiers

source 1..4676

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match	68.2%;	Score 2128;	DB 2;	Length 4676;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 2178;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	AAGTCTAAAGCCTCCGAGGCCAAGCGCGTGTCTACTGCGCGCGTCTTCTTAGTGGCG	60
DB	1	AAGTCTAAAGCCTCCGAGGCCAAGCGCGTGTCTACTGCGCGCGTCTTCTTAGTGGCG	60
QY	61	CGTTGCGCGCGTGGTGTCTACCGCGCGCGCGCGCGAGGAGCACTGCACACGAGACCG	120
DB	61	CGTTGCGCGCGTGGTGTCTACCGCGCGCGCGCGCGAGGAGCACTGCACACGAGACCG	120
QY	121	GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATGAGCGCAGTGGGACGG	180
DB	121	GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATGAGCGCAGTGGGACGG	180
QY	181	TCCGCGCGCGCGCGCGGAGGAGAAATGCAACATGCGCAGCAGTGGAAACAGAACAG	240
DB	181	TCCGCGCGCGCGCGCGGAGGAGAAATGCAACATGCGCAGCAGTGGAAACAGAACAG	240
QY	241	CTGGGTTGAGATATTTGAACTCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG	300
DB	241	CTGGGTTGAGATATTTGAACTCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG	300
QY	301	CCTAAATTGAGCGCTTTTATGTTGAGCGGTTATCTGAGGATGAGCTTAAAGAGTGCCT	360
DB	301	CCTAAATTGAGCGCTTTTATGTTGAGCGGTTATCTGAGGATGAGCTTAAAGAGTGCCT	360
QY	361	GCGGATACAGAAATATCATGGCTACATGATGGCTAAGGACCAACATGATTTTCACTGTT	420
DB	361	GCGGATACAGAAATATCATGGCTACATGATGGCTAAGGACCAACATGATTTTCACTGTT	420
QY	421	GTGAGAGGAATGATCCAGATGAGCTCAATTCAGACAGAACTTATACCTTGGCATGTCT	480
DB	421	GTGAGAGGAATGATCCAGATGAGCTCAATTCAGACAGAACTTATACCTTGGCATGTCT	480
QY	481	GGTGAGAACAGAGAAATACATGTTTATTTCTGAAATTCGAAATCCGAAACTATCAATAGAGCA	540
DB	481	GGTGAGAACAGAGAAATACATGTTTATTTCTGAAATTCGAAATTCGAAACTATCAATAGAGCA	540
QY	541	GCAGTCTTAATGCTCTCTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTTGGACTAT	600
DB	541	GCAGTCTTAATGCTCTCTTGGAGCCTCTTTTGGATCTTTTTCAGGCAACACTTGGACTAT	600
QY	601	GGAATGTAATCTCCGAGAGAGAACTATTAAGAGAGAAAGAAACGCAATGGAAACAGTCGGA	660
DB	601	GGAATGTAATCTCCGAGAGAGAACTATTAAGAGAGAGAAAGAAACGCAATGGAAACAGTCGGA	660
QY	661	ATTGCTTCTTACGATATACCAAGGAGTGGACATTTCTGTTTCAAGCCGGTAGTGA	720
DB	661	ATTGCTTCTTACGATATACCAAGGAGTGGACATTTCTGTTTCAAGCCGGTAGTGA	720
QY	721	ATTATACAGTAAAGATGAGGCGCCACAAGGATTTACGCAACACTTTTAAGGCCAAT	780
DB	721	ATTATACAGTAAAGATGAGGCGCCACAAGGATTTACGCAACACTTTTAAGGCCAAT	780
QY	781	CTAGTGGAAACTAGTGTGCCAACAATACGAGATGATCCAAATTTATGCCCGCTGATCCA	840
DB	781	CTAGTGGAAACTAGTGTGCCAACAATACGAGATGATCCAAATTTATGCCCGCTGATCCA	840

QY	841	GACTGGATTCCTTTTATACATAGCAAGCATATTTGGATATCTAAACATCGTAACCAAGAGAA	900
DB	841	GACTGGATTCCTTTTATACATAGCAAGCATATTTGGATATCTAAACATCGTAACCAAGAGAA	900
QY	901	GAAAGGAGACTCACTTATGTGCAATAGCTAGCCCAACATGGAAGAGATGCCAGATCA	960
DB	901	GAAAGGAGACTCACTTATGTGCAATAGCTAGCCCAACATGGAAGAGATGCCAGATCA	960
QY	961	GCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
DB	961	GCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
QY	1021	TGTCCTAAAGAGCTGAAACCAACTCCCAAGTGGTGGTAAATTTCTTAGAATTTCTATAGAGAA	1080
DB	1021	TGTCCTAAAGAGCTGAAACCAACTCCCAAGTGGTGGTAAATTTCTTAGAATTTCTATAGAGAA	1080
QY	1081	AATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAAACCAAGGAGG	1140
DB	1081	AATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAAACCAAGGAGG	1140
QY	1141	GCAGATTCATTCGGTTATCTTAAACAGGTACAGCAAACTCTTAAAGTCACTTTTAAAGATG	1200
DB	1141	GCAGATTCATTCGGTTATCTTAAACAGGTACAGCAAACTCTTAAAGTCACTTTTAAAGATG	1200
QY	1201	TCAGAAATTAATGATGCTGGAAGAGGATCATAGATGTCTAGATAGAGAACTAATTT	1260
DB	1201	TCAGAAATTAATGATGCTGGAAGAGGATCATAGATGTCTAGATAGAGAACTAATTT	1260
QY	1261	CAACCTTTTGAATTCATTTTCAAGGAGTGAATAATATTTGCCAGAGCTGGATGACTCCT	1320
DB	1261	CAACCTTTTGAATTCATTTTCAAGGAGTGAATAATATTTGCCAGAGCTGGATGACTCCT	1320
QY	1321	GAGGAAATATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCTACAGATAGTG	1380
DB	1321	GAGGAAATATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCTACAGATAGTG	1380
QY	1381	TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATTTATGGAAGGAGGAGGAGCTC	1440
DB	1381	TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATTTATGGAAGGAGGAGGAGCTC	1440
QY	1441	ATTGAGTCAGTGCCTGATTTCTGAGCGCACTAAATTTCTATGAAGAAACCAACAGACATC	1500
DB	1441	ATTGAGTCAGTGCCTGATTTCTGAGCGCACTAAATTTCTATGAAGAAACCAACAGACATC	1500
QY	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGTACGAGAGGAAATTCAG	1560
DB	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGTACGAGAGGAAATTCAG	1560
QY	1561	TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTTTATACAAAATTACATCTATT	1620
DB	1561	TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTTTATACAAAATTACATCTATT	1620
QY	1621	TTTAAAGAAAGCAAAATATAAATGATGCTGGTGGGCTGCTCTCCAAAGTATTTCAG	1680
DB	1621	TTTAAAGAAAGCAAAATATAAATGATGCTGGTGGGCTGCTCTCCAAAGTATTTCAG	1680
QY	1681	TGTCCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGCGAAGTCTTGGCCGGCAT	1740
DB	1681	TGTCCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGCGAAGTCTTGGCCGGCAT	1740
QY	1741	GGATCTTAATATCCAAGTTGATGAGTCAAGGCTGGTATATTTTGAAGGACCAAGAGAC	1800
DB	1741	GGATCTTAATATCCAAGTTGATGAGTCAAGGCTGGTATATTTTGAAGGACCAAGAGAC	1800
QY	1801	TCCCTTTTAGAGCATCACTGTACGTAGTCACTTAATCTGAGAGGAGTGAACAGG	1860
DB	1801	TCCCTTTTAGAGCATCACTGTACGTAGTCACTTAATCTGAGAGGAGTGAACAGG	1860
QY	1861	CTGACTGACCGTGGCTACTCATTCTTGTGTCATCAGTACGACTGTGACTCTTTTATA	1920
DB	1861	CTGACTGACCGTGGCTACTCATTCTTGTGTCATCAGTACGACTGTGACTCTTTTATA	1920

QY	1921	AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAAGTATCAAGTCTT	1980
Db	1921	AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAAGTATCAAGTCTT	1980
QY	1981	GAAAGATGACCCCAACTTTCAGAAACAAAGAAATTTGGGCGACAGATTTGGATTACAGAGGT	2040
Db	1981	GAAAGATGACCCCAACTTTCAGAAACAAAGAAATTTGGGCGACAGATTTGGATTACAGAGGT	2040
QY	2041	CCTCTCTCTGACTACTCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACA	2100
Db	2041	CCTCTCTCTGACTACTCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACA	2100
QY	2101	TTGTATGGAGTCTCTACAGGCTCATGATCTACAGCTCGGAAAGAAATATCTACTGTG	2160
Db	2101	TTGTATGGAGTCTCTACAGGCTCATGATCTACAGCTCGGAAAGAAATATCTACTGTG	2160
QY	2161	CTGTTCATATATGTTGTC	2179
Db	2161	CTGTTCATATATGTTGTC	2179
RESULT 18			
AX405770			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Tang, Y. T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, P.,			
Xue, A. J., Yang, Y., Wehrman, T. and Drmanac, R. T.			
Novel nucleic acids and polypeptides			
Patent: WO 0222660-A 185 21-MAR-2002;			
HYSEQ, INC. (US)			
FEATURES			
Source			
1..2842			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
234..2582			
/note="unnamed protein product"			
/codon_start=1			
/protein_id="CAD34877.1"			
/db_xref="GI:21439980"			
/translation="MAAAWTEQLGVIEPETAOCERNIESQDRPKLEPFYVERYSWSQ			
LKKLADPKHGMMAKAPHDPMFKRNDPDPHSDRIYYLAMSGENRENTLFYSEI			
PKTINRAAYLMSWPKLDLDFQATLDYGMYSREBELLRKKRIGTVGIASDYDHQSG			
TFLFGAGSGITHVKGQPGQFTQQLRPNLVETSCPNIRMDPKLPADPDWIAPIHNS			
DIWISNIVREERLTYHNELANMEEDARSAGVATFVLQEBFDYRSGYVWCPKAETI			
DSGKILRIEENDESEVEIIRVTSPLMRETRADSFYKPTGTANPKVTFKMSIMI			
DARGKIIDIDKELIQPFILFEGVEYIARAGWTPGKYAWSILLDRSQRLQIWLIS			
FELFIPVEDDVMERLIESVPSVTPLIITYEETDWINIHDIHFVFPQSEHEIEIF			
IFASECKTFGRHLKYITSLIKSKYRSGLPAPSDFKCPIKEIAITSGSEWEVLGR			
HGSNIQDVNRRLVYFPGTKDSLPLEHLLVYVSVNPFVETRLDRGYSKSCISQHCD			
FPIYSKNQKPHCVSLYKLSPEDDPTCKTEFWATILDSAGPLPDVTPPEIFSPES			
TGTFLYGLMKVHDLQPKKYPVLFYFGQPOVIAAGAPVLMIFDYDGYERYMVG			
PDQNEQGYLGSVMQAEKFPSPBNRLLHLHGFLDENVHFAHTSILLSFLVRAGKPYD			
LOIYPOERHSIRVPSGSEHYELHLLHYLOENLGSRLAALKVI"			
ORIGIN			
Query Match			
Best Local Similarity 67.8%; Score 2116; DB 2; Length 2842;			
Matches 2166; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY			
20 GCCAAGCCGCTGCTACTGCGCGCTGCTTCTTAGTGGCGGCTTGC			
Db			
40 GCCAAGCCGCTGCTACTGCGCGCTGCTTCTTAGTGGCGGCTTGC			

QY	80	CACCGCGCGCGCGCGAGGAAGCCACCTGAGGACCGAGTGTGAGGCGGCGAGCA	139
Db	100	CACCGCGCGCGCGCGAGGAAGCCACCTGAGGACCGAGTGTGAGGCGGCGAGCA	159
QY	140	TGAAGCGGCGCGAGGCGCGCTCCATAGGCGCGTCCGAGCGGCGGCGGCGG	199
Db	160	TGAAGCGGCGCGAGGCGCGCTCCATAGGCGCGTCCGAGCGGCGGCGGCGG	219
QY	200	AAGGAAATGCAACATGCGCAGCAGCAATGGAACACAGACAGCTGGGTGTGAGATATTG	259
Db	220	AAGGAAATGCAACATGCGCAGCAGCAATGGAACACAGACAGCTGGGTGTGAGATATTG	279
QY	260	AAACTGCGGACTGTGAGGAGATATTGAATCAGAGATCGGCGCTAAATTTGGAGCCTTTT	319
Db	280	AAACTGCGGACTGTGAGGAGATATTGAATCAGAGATCGGCGCTAAATTTGGAGCCTTTT	339
QY	320	ATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGTTGCCGATACAGAAAAATATC	379
Db	340	ATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGTTGCCGATACAGAAAAATATC	399
QY	380	ATGGCTACATGATGGCTAAGGCACCATCATGATTTTCATGTTGTGAAGAGGAATGATCCAG	439
Db	400	ATGGCTACATGATGGCTAAGGCACCATCATGATTTTCATGTTGTGAAGAGGAATGATCCAG	459
QY	440	ATGGAACCTCATTCAGACAGAAATCTATTACCTTGCCCATGTCTGGTGAAGACAGAGAAATA	499
Db	460	ATGGAACCTCATTCAGACAGAAATCTATTACCTTGCCCATGTCTGGTGAAGACAGAGAAATA	519
QY	500	CACCTGTTTATTTCTGAAATTTCCAAAACTATCAATAGAGCAGCAGTCTTAAATGCTCTCT	559
Db	520	CACCTGTTTATTTCTGAAATTTCCAAAACTATCAATAGAGCAGCAGTCTTAAATGCTCTCT	579
QY	560	GGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGACATATGGAATGATTTCTCGAGAAG	619
Db	580	GGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGACATATGGAATGATTTCTCGAGAAG	639
QY	620	AAGAACTATTAAAGAGAAAGAAACGCAATTTGGAACAGTGGAAATTCCTTCTACGATATC	679
Db	640	AAGAACTATTAAAGAGAAAGAAACGCAATTTGGAACAGTGGAAATTCCTTCTACGATATC	699
QY	680	ACCAAGAGAGTGGAAATTTCTGTTTCAAGCGCGTAGTGGAAATTTATCACTTAAAGATG	739
Db	700	ACCAAGAGAGTGGAAATTTCTGTTTCAAGCGCGTAGTGGAAATTTATCACTTAAAGATG	759
QY	740	GAGGGCCACAGGATTTACGCAACACCTTTAAAGGCCCAATCTAGTGGAAATCTAGTTGTC	799
Db	760	GAGGGCCACAGGATTTACGCAACACCTTTAAAGGCCCAATCTAGTGGAAATCTAGTTGTC	819
QY	800	CCAACATACGATGGATCCAAATTTATGCGCGCTGATCCAGCTGGATGCTTTTATAC	859
Db	820	CCAACATACGATGGATCCAAATTTATGCGCGCTGATCCAGCTGGATGCTTTTATAC	879
QY	860	ATAGCAACGATTTTGGATATCTAACATCGTAAACAGAGAAAGGAGACTCACTTATG	919
Db	880	ATAGCAACGATTTTGGATATCTAACATCGTAAACAGAGAAAGGAGACTCACTTATG	939
QY	920	TGCAGATACGCTACCCACATGGAAGAGATGCGAGATCAGCTGGAGTGCCTACCTTTG	979
Db	940	TGCAGATACGCTACCCACATGGAAGAGATGCGAGATCAGCTGGAGTGCCTACCTTTG	999
QY	980	TTCTCCAGAGAAATTTGATAGATATTCTGCTATTGTTGGTGTCCAAAAGCTGAAACAA	1039
Db	1000	TTCTCCAGAGAAATTTGATAGATATTCTGCTATTGTTGGTGTCCAAAAGCTGAAACAA	1059
QY	1040	CTCCAGTGGTGGTAAAAATTTTGAATTTCTATATGAAGAAAAATGATGAATCTGAGGTGG	1099
Db	1060	CTCCAGTGGTGGTAAAAATTTTGAATTTCTATATGAAGAAAAATGATGAATCTGAGGTGG	1119
QY	1100	AAATTTATTCATGTTACATCCCTTATGTTGGAACAGAGGCGGAGATTCAATCCGTTATC	1159
Db	1120	AAATTTATTCATGTTACATCCCTTATGTTGGAACAGAGGCGGAGATTCAATCCGTTATC	1179
QY	1160	CTAAAAAGGTAACAGAAATCTTAAAGTCACTTTTAAAGTGTGTCAGAAAAATATGATTGATG	1219

Db	1180	CTAAGACGGTACAGCAATCCTTAAGTCACTTTTAAAGTGCAGAAATAATGATTGATG	1239
Qy	1220	CTGAAGAGGATCATAGATGTCATAGATGAAGAACTAATTCACCTTTTGAGATTCAT	1279
Db	1240	CTGAAGGAGGATCATAGATGTCATAGATGAAGAACTAATTCACCTTTTGAGATTCAT	1299
Qy	1280	TGAAGAGGTTGATATATATGACAGAGCTGAGTCTCTGAGGAAATAATGCTTGGT	1339
Db	1300	TGAAGGAGTTGATATATATGACAGAGCTGAGTCTCTGAGGAAATAATGCTTGGT	1359
Qy	1340	CCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATAT	1399
Db	1360	CACTCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATAT	1419
Qy	1400	TTATCCCACTAGAAGATGATTTATGAAAGGAGAGACTCATTCAGTCAGTCCCTGATT	1459
Db	1420	TTATCCCACTAGAAGATGATTTATGAAAGGAGAGACTCATTCAGTCAGTCCCTGATT	1479
Qy	1460	CTGTGACGCGCACTAATTCATCTATGAAGAAACACAGACATCTCGATAAATATCCATGACA	1519
Db	1480	CTGTGACGCGCACTAATTCATCTATGAAGAAACACAGACATCTCGATAAATATCCATGACA	1539
Qy	1520	TCCTTCATGTTTTTCCCAAGTCAAGAGGAGAAATGAGTTTATTTTTCCTCTGAT	1579
Db	1540	CTCTTCATGTTTTTCCCAAGTCAAGAGGAGAAATGAGTTTATTTTTCCTCTGAT	1599
Qy	1580	GCAAAACAGGTTTCCTGTCATTTATACAAATTCATCTATTTAAAGGAAAGCAATATATA	1639
Db	1600	GCAAAACAGGTTTCCTGTCATTTATACAAATTCATCTATTTAAAGGAAAGCAATATATA	1659
Qy	1640	AACGATCCAGTGGTGGTCTGCTCAAGTGAATTCAGTGTCTCTATCAAGAGGACA	1699
Db	1660	AACGATCCAGTGGTGGTCTGCTCAAGTGAATTCAGTGTCTCTATCAAGAGGACA	1719
Qy	1700	TAGCAATACAGTGTGAATGGGAAGTTCTTGCCGCGCATGATCTAATATCAAGTTG	1759
Db	1720	TAGCAATACAGTGTGAATGGGAAGTTCTTGCCGCGCATGATCTAATATCAAGTTG	1779
Qy	1760	ATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCAC	1819
Db	1780	ATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCAC	1839
Qy	1820	TGTCAGTGTGAGTACGTAATCTGAGAGGTGAAGGCTGACTGACCGTGGCTACT	1879
Db	1840	TGTCAGTGTGAGTACGTAATCTGAGAGGTGAAGGCTGACTGACCGTGGCTACT	1899
Qy	1880	CACATTCCTGCTGATCAGTCAAGTGTGCTTTTAAAGTAAGTATAGTAACGAGA	1939
Db	1900	CACATTCCTGCTGATCAGTCAAGTGTGCTTTTAAAGTAAGTATAGTAACGAGA	1959
Qy	1940	AGAATCCACACTGTGTCCCTTTTACAAGCTATCAAGTCTGAGAGTACCCAACTTGCA	1999
Db	1960	AGAATCCACACTGTGTCCCTTTTACAAGCTATCAAGTCTGAGAGTACCCAACTTGCA	2019
Qy	2000	AAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGAGGTCCTCTTCTGACTATACT	2059
Db	2020	AAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGAGGTCCTCTTCTGACTATACT	2079
Qy	2060	CTCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTTACA	2119
Db	2080	CTCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTTACA	2139
Qy	2120	AGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGCTC	2179
Db	2140	AGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGCTC	2199
Qy	2180	CTCAGGT 2186	
Db	2200	CTCAGGT 2206	

BC030688
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
CONSTRM
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT

BC030688 3125 bp mRNA linear PRI 19-JAN-2006
Homo sapiens dipeptidyl-peptidase 8, transcript variant 4, mRNA
(cDNA clone MGC:26191 IMAGE:4822550), complete cds.
BC030688
BC030688.2 GI:34190028
MGC.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 3125)
Straussberg,R.D., Feingold,B.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altshuler,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heideh,F.,
Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S.S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,B.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3125)
NIH MGC Project
Direct Submission
Submitted (24-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
On Aug 25, 2003 this sequence version replaced gi:21265132.
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadansystemsbiology.org
Anup Madan, Jessica Fahney, Erin Helton, Mark Kerteman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 33 Row: d Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 37577090.
Location/Qualifiers
1. 3125
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:26191 IMAGE:4822550"
/tissue_type="Testis"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescriptR"
1. 3125
/gene="DPP8"

FEATURES
source
gene

/note="synonyms: FLJ20283, FLJ14920, MGC26191, MSTP141, DP8, DRP1",			
/db_xref="GeneID:54878"			
/db_xref="HGNC:16490"			
/db_xref="IMGT/GENE-DB:16490"			
/db_xref="MIM:606819"			
319..2862			
/gene="DPP8"			
/codon_start=1			
/product="dipeptidyl peptidase 8, isoform 4"			
/protein_id="AAH30688.3"			
/db_xref="GI:54114905"			
/db_xref="GeneID:54878"			
/db_xref="HGNC:16490"			
/db_xref="IMGT/GENE-DB:16490"			
/db_xref="MIM:606819"			
/translations="MMKRSQMKIKSGKCNMAAAMETBQLGVEIFETADCEENIESOD RPLKPPYVERVSWQLKLLADTKYHGMKAPDFMFKVKNRNDPDGPHSLRIYVL AMSGENRNTLFYSEIPKTIINRAAVLMLSWKPLLDFOATLDYGMVSRREBELRERK IGTVGASVDYHOGSGTFLFQAGSGIYHVKGQPGQPTQPLRPNLVTCPNIRWDP KLCPADPDWIAFTHSNDIWI SINIVTRERRLLTYVHNELANNEEDSAGVATFVLOEE FDRYGVWCPKAEITTPSGKILRLIYBENDESEVEIIHVTSPMLERRADSPRYKWS GTANPKVTFKMSIEMIDAGRIIDVIDKELIQPEILLFEGVEYLARAGWTPREGKWS ILLDRSOTRLOI VLISPELFIPEVDVNERORLIESVPDSVTPILIIYEETDIIWIIH DIHFVPQSHSEBEIEFIPASECKTGFRLHKITISLKESKYKSSSGLPAPSPKPI KERIATISGEWELGRHGSINOVRRLVYFECTKQSPLEHLXYVSYVNRGEVEL TDGYSHSCCISQHOCPFTISKYNKPNHCVSLYKLSPPEDDPTCKTKERFWALISA GPLPDDTTPPEIFSFSTGTFLYGLMLYKPHDLQPKYPTVLFIYGGPQVLVNRFK GVKYPRLNTLASLGYYVAVVDNRGSHRGLKFEKAPKYKMAIAGAPVILNIWIFIDGY TERYMGPDNQGGYVLAGSVAQAEKPPSENRULLLHGFLDENVHFHTSILLISFLV RAGKPYDLQIYQPERHSIRVPESGHEHYELHLHLTLQENLGSRIANLKV1"			
ORIGIN			
Query Match		66.7%; Score 2081; DB 5; Length 3125;	
Best Local		Similarity 99.9%; Pred. No. 0;	
Matches 213;		Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
Qy	201	AGGAAATGCAATCGGAGCGAGCAATGGAACAGAACAGACAGCTGGGTGTGAGATATTGA	260
Db	354	AGGAAATGCAATCGGAGCGAGCAATGGAACAGAACAGACAGCTGGGTGTGAGATATTGA	413
Qy	261	AATCGCGACTGTGAGGAGAAATTTGAATACAGGATCGGCTAAATTTGGAGCTTTTGA	320
Db	414	AATCGCGACTGTGAGGAGAAATTTGAATACAGGATCGGCTAAATTTGGAGCTTTTGA	473
Qy	321	TGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACAGAAATATCA	380
Db	474	TGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACAGAAATATCA	533
Qy	381	TGCTACATGATGGCTAAGGACACACATGATTTTATGTTTGTGAAGAGGAATGATCAGA	440
Db	534	TGCTACATGATGGCTAAGGACACACATGATTTTATGTTTGTGAAGAGGAATGATCAGA	593
Qy	441	TGGACCTCATTCAGACAGAAATCTATTACTTGGCCATGCTGGTGAGAAACAGAGAAATAC	500
Db	594	TGGACCTCATTCAGACAGAAATCTATTACTTGGCCATGCTGGTGAGAAACAGAGAAATAC	653
Qy	501	ACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCAGCAGTCTTATGCTCTCTTG	560
Db	654	ACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCAGCAGTCTTATGCTCTCTTG	713
Qy	561	GAAGCCTCTTTTGGATCTTTTTCAGGCAACACTCGACTNTGGATGATTTCTCGAGAAGA	620
Db	714	GAAGCCTCTTTTGGATCTTTTTCAGGCAACACTCGACTNTGGATGATTTCTCGAGAAGA	773
Qy	621	AGAACTATTAAAGAAAGAAACGCAATGGAACTGGAACTGGCTTCTTACGATTATCA	680
Db	774	AGAACTATTAAAGAAAGAAACGCAATGGAACTGGAACTGGCTTCTTACGATTATCA	833
Qy	681	CCAAAGAGTGGAAATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACTGTAAGAGATGG	740
Db	834	CCAAAGAGTGGAAATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACTGTAAGAGATGG	893

Qy	741	AGGGCCACAGAGATTTAGCCAAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCC	800
Db	894	AGGGCCACAGAGATTTAGCCAAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCC	953
Qy	801	CAACATACGGATGGATCCAAATTTATCCCGCTGATCCAGACTGAGATTGCTTTTATACA	860
Db	954	CAACATACGGATGGATCCAAATTTATCCCGCTGATCCAGACTGAGATTGCTTTTATACA	1013
Qy	861	TAGCAACGATATTTGGATATCTAACATCGTAACAGAGAGAGAGAGACTCACTATGCT	920
Db	1014	TAGCAACGATATTTGGATATCTAACATCGTAACAGAGAGAGAGAGACTCACTATGCT	1073
Qy	921	GCACATGAGCTAGCCAACTGGAAGAGATCCAGATCAGCTGGAGTGGCTACCTTTGT	980
Db	1074	GCACATGAGCTAGCCAACTGGAAGAGATCCAGATCAGCTGGAGTGGCTACCTTTGT	1133
Qy	981	TCTCCAGAGAGATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAGAGCTGAAACAAC	1040
Db	1134	TCTCCAGAGAGATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAGAGCTGAAACAAC	1193
Qy	1041	TCCAGTGGTGGTAAATTTCTTGAATCTTATATGAAGAAATGATGAATCTGAGGTGA	1100
Db	1194	TCCAGTGGTGGTAAATTTCTTGAATCTTATATGAAGAAATGATGAATCTGAGGTGA	1253
Qy	1101	AATTAATCATGTTACATCCCCTATGTTGGAACAAGAGAGGAGAGATTCATTCGGTATCC	1160
Db	1254	AATTAATCATGTTACATCCCCTATGTTGGAACAAGAGAGGAGAGATTCATTCGGTATCC	1313
Qy	1161	TAAACAGGTAACAGCAATCCTTAAAGTCACTTTTAAGATGTGCAGAAATTAATGATTGATGC	1220
Db	1314	TAAACAGGTAACAGCAATCCTTAAAGTCACTTTTAAGATGTGCAGAAATTAATGATTGATGC	1373
Qy	1221	TGAAGGAGGATCATAGATGTCTATAGATAAGGAACTAATTTCAACCTTTTGAGATTCTATT	1280
Db	1374	TGAAGGAGGATCATAGATGTCTATAGATAAGGAACTAATTTCAACCTTTTGAGATTCTATT	1433
Qy	1281	TGAAGGAGTTCGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGAGAAATATGCTTGTC	1340
Db	1434	TGAAGGAGTTCGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGAGAAATATGCTTGTC	1493
Qy	1341	CATCCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTATT	1400
Db	1494	CATCCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTATT	1553
Qy	1401	TATCCAGTAAAGATGATGTTATGGAAGGAGAGAGACTCACTTGAGTCAGTGCCTGATTC	1460
Db	1554	TATCCAGTAAAGATGATGTTATGGAAGGAGAGAGACTCACTTGAGTCAGTGCCTGATTC	1613
Qy	1461	TGTGACGCCACTAATTTATCTATGAAGAAACACACACATCTGGATAAATATCCATGACAT	1520
Db	1614	TGTGACGCCACTAATTTATCTATGAAGAAACACACACATCTGGATAAATATCCATGACAT	1673
Qy	1521	CTTTCTATGTTTTCCTCCAAAGTCAAGAGAGAAATTCAGTTATTTTGGCTCTGAAATG	1580
Db	1674	CTTTCTATGTTTTCCTCCAAAGTCAAGAGAGAAATTCAGTTATTTTGGCTCTGAAATG	1733
Qy	1581	CAAAACAGTTTTCCTGCTCATTTATCAAAATTCATCTATTTTAAAGGAAAGCAATATPAA	1640
Db	1734	CAAAACAGTTTTCCTGCTCATTTATCAAAATTCATCTATTTTAAAGGAAAGCAATATPAA	1793
Qy	1641	ACGATCAGTGTGGGCTGCTGCTCCCAAGTATTTCAAGTATTTCAAGTGTCTCTATCAAGAGGAGAT	1700
Db	1794	ACGATCAGTGTGGGCTGCTGCTCCCAAGTATTTCAAGTATTTCAAGTGTCTCTATCAAGAGGAGAT	1853
Qy	1701	AGCAATTAACAGTGTGAAATGGAAATGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGA	1760
Db	1854	AGCAATTAACAGTGTGAAATGGAAATGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGA	1913
Qy	1761	TGAAGTCAAGAGCTGGTATATTTTGAAGGACCAAAAGCTCCCTCTTTAGAGCATCACCT	1820
Db	1914	TGAAGTCAAGAGCTGGTATATTTTGAAGGACCAAAAGCTCCCTCTTTAGAGCATCACCT	1973
Qy	1821	GTACGTAGTCAGTTACGTTAAATCCTGAGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTC	1880

Db	1974	GTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTC	2033
Qy	1881	ACATTCCTTGCTGCATCAGTCAGCACTGTGACTCTCTTATTAAGTAAGTATAGTAACAGAA	1940
Db	2034	ACATTCCTTGCTGCATCAGTCAGCACTGTGACTCTCTTATTAAGTAAGTATAGTAACAGAA	2093
Qy	1941	GAATCCACACTGTGTGCTCCCTTTTAAAGCTATCAAGTCTCTGAAGATGAACCACTTGC	2000
Db	2094	GAATCCACACTGTGTGCTCCCTTTTAAAGCTATCAAGTCTCTGAAGATGAACCACTTGC	2153
Qy	2001	AACAAAGAAATTTTGGGCCCAATTTTGGATTTCAGCAGGTCCTCTCTGACTATACTCC	2060
Db	2154	AACAAAGAAATTTTGGGCCCAATTTTGGATTTCAGCAGGTCCTCTCTGACTATACTCC	2213
Qy	2061	TCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAA	2120
Db	2214	TCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAA	2273
Qy	2121	GCCTCATGATCTACAGCCTGGAAGAAATATCTCTGCTGCTGTTTATATATGGTGTCC	2180
Db	2274	GCCTCATGATCTACAGCCTGGAAGAAATATCTCTGCTGCTGTTTATATATGGTGTCC	2333
Qy	2181	TCAGGTGAGTGTGGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCT	2240
Db	2334	TCAGGTGAGTGTGGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCT	2393
Qy	2241	AGCCTCTCTAGGTTATGTTGTTAGTAGATAGACACAGGGATCTCTGTCACCGAGGCT	2300
Db	2394	AGCCTCTCTAGGTTATGTTGTTAGTAGATAGACACAGGGATCTCTGTCACCGAGGCT	2453
Qy	2301	TAAATTTGAAGCGCTTTTAAATATAAAATGG	2332
Db	2454	TAAATTTGAAGCGCTTTTAAATATAAAATGG	2485
RESULT 20			
AR631288			
LOCUS			
Sequence 22 from patent US 6844180.			
AR631288			
ACCESSION			
AR631288.1			
VERSION			
GI:59770936			
KEYWORDS			
Unknown.			
SOURCE			
Unknown.			
ORGANISM			
Unclassified.			
1 (bases 1 to 4685)			
REFERENCE			
O.S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.			
Serrine protease genes related to DppIV			
Patent: US 6844180-A 22 18-JAN-2005;			
Ferring BV,;			
NLX;			
FEATURES			
Location/Qualifiers			
1..4685			
/organism="unknown"			
/mol_type="genomic DNA"			
ORIGIN			
Query Match			
Best Local Similarity 63.6%; Score 1984; DB 2; Length 4685;			
Matches 2034; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	AAGTGCTAAAGCTCCGAGGCCAAGCGCTGCTACTGCGCGCTGCTTCTAGTGCGG	60
Db	1	AAGTGCTAAAGCTCCGAGGCCAAGCGCTGCTACTGCGCGCTGCTTCTAGTGCGG	60
Qy	61	CGTTCCGCGCTGGGTTGTACCGCGCCCGCCGCGGAGGACCTGCAACCGAGCCG	120
Db	61	CGTTCCGCGCTGGGTTGTACCGCGCCCGCCGCGGAGGACCTGCAACCGAGCCG	120
Qy	121	GAGTGGAGCGCGCAGCATGAAGCGCGCAGCGCGCTCATAGCGCAGTCTGGAGCGG	180
Db	121	GAGTGGAGCGCGCAGCATGAAGCGCGCGCGCGCTCATAGCGCAGTCTGGAGCGG	180

QY 1261 CAACCTTTTGAGATCTTATTTGAGGAGTTGCAATATATTTGCCAGAGCTGGATGACTCT 1320
Db |||||
QY 1321 GAGGGAAATATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCTCCATACAGATAGTG 1380
Db |||||
QY 1321 GAGGGAAATATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCTCCATACAGATAGTG 1380
Db |||||
QY 1381 TTGATCTCACCTCGAATTTATTTATCCAGTAGAGATGATGTTATGAAAAGGAGAGACTC 1440
Db |||||
QY 1441 ATTGAGTCAGTCCCTGATTTCTGTGACGCCACTAATATCTATGAAGAAACAACAGACATC 1500
Db |||||
QY 1501 TGGATAATATCCATGATCATCTTTTCATGTTTTTCCCAAGTCACGAGAGGAAATTCAG 1560
Db |||||
QY 1561 TTTATTTTGGCTCTGAATCAAAACAGAGTTTCCCGTCAATTTATACAAAATTTACATCTATT 1620
Db |||||
QY 1621 TTAAGGAAAGCAAAATATAAACGATCCAGTGGTGGCTCGCTCCAGTGATTTCAAG 1680
Db |||||
QY 1681 TGCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGAGTTCTTGGCGGCAT 1740
Db |||||
QY 1741 GGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCCAAGAC 1800
Db |||||
QY 1801 TCCCTTTAGAGCATCACCTGTACGTAGTCAGTACGTAAATCTCGAGAGGTGACAAG 1860
Db |||||
QY 1861 CTGACTGACCGTGGTACTCACATCTTGTGTCATCAGTCAGTCAGCTGTGACTTCTTTATA 1920
Db |||||
QY 1921 AGTAAGTATAGTAACAGAGAAATCCACATCTGTGTGCTTTCAGTACGACTGTGACTTCTTTATA 1920
Db |||||
QY 1921 AGTAAGTATAGTAACAGAGAAATCCACATCTGTGTGCTTTCAGTACGACTGTGACTTCTTTATA 1920
Db |||||
QY 1981 GAAGATGACCCACTTGCANAAAGGAATTTTGGGCCACCAATTTTGGATTTCAG 2035
Db |||||

RESULT 21
AX608745
LOCUS AX608745
DEFINITION Sequence 22 from Patent WO0231134.
ACCESSION AX608745
VERSION AX608745.1
KEYWORDS GI:28404308
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Qi,S., Akinsanya,K.O., Riviere,P.J. and Junien,J.L.
TITLE Novel serine protease genes related to dppv
JOURNAL Patent: WO 0231134-A 22 18-APR-2002;
Ferring BV (NL)
FEATURES
Location/Qualifiers
1..4685
/organism="Homo sapiens"

Query Match 63.6%; Score 1984; DB 2; Length 4685;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2034; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
ORIGIN
QY 1 AGTGTCTAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTTCTTTAGTGCCG 60
Db |||||
QY 61 CGTTCCGCGCTGGGTTTGTCAACCAGCCGCGCGCGAGGAAGCACTGCAACAGGACCG 120
Db |||||
QY 121 GAGTGGAGCGCGCGAGCATGAAAGCGCGCAGAGCCCGCTCCATAGCGCAGCTCGGACCG 180
Db |||||
QY 181 TCCGCGCGCGCGCGGGGGAAGGAAATGCAACATGCGCAGCAGCAATGGAACAGAACAG 240
Db |||||
QY 241 CTGGGTGTGAGATATTTGAAACTGCGGGACTGTGAGGAGAAATTTGAAATCACAGGATCG 300
Db |||||
QY 301 CCTAAATTGGAGCCCTTTTATTTGAGCGGTAATTCCTGGAGTCAGCTTAAAGCTGCTT 360
Db |||||
QY 361 GCCGATACCAAGAAATATCATGGCTACATGATGGCTTAAGCAGCACCATGATTTTCTTT 420
Db |||||
QY 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTTATTTACCTTGGCATGCT 480
Db |||||
QY 481 GGTGAGAAACAGAAATAACACTGTTTATTTCTGAAATCCCAAACTATCAATAGAGCA 540
Db |||||
QY 541 GCAGTCTTAATGCTCTCTTGGAGCCTCTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db |||||
QY 601 GGAATGTATTCTCGAAGAGAACTNTTAAGAGAAAGAAACGCAATGGAAACAGTCGGA 660
Db |||||
QY 661 ATTGCTTCTTACGATTTATCCCAAGAAAGCAATTTCTGTTTCAAGCCGCTAGTGGG 720
Db |||||
QY 721 ATTATACAGTAAAGATGGAGGGCCACAAGATTTACGCAACACCTTTTAAGGCCCAAT 780
Db |||||
QY 781 CTAGTGAAACTAGTGTGCCAACATACGATGGAATCCAAATTTATGCCCTGCTGATCCA 840
Db |||||
QY 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCTGTAACAGAG 900
Db |||||
QY 901 GAAAGAGACTCATCTTATGTGCACAANTGAGCTAGCCAAACATGGAAGAAAGATGCCAGATCA 960
Db |||||
QY 961 GCTGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGCTATTGCTG 1020

/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Db 121 TCGGACGGTCCGGCGGGCCGGGGAAGGAAATCCACATGGCAGCAGCAATGGAA 180
Qy 232 ACAGAAACAGCTGGGTGTGAGATATTTGAACTCGGACCTGTGAGGAGATATTTGAATCA 291
Db 181 ACAGAAACAGCTGGGTGTGAGATATTTGAACTCGGACCTGTGAGGAGATATTTGAATCA 240
Qy 292 CAGGATCGGCCTAAATTCGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAA 351
Db 241 CAGGATCGGCCTAAATTCGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAA 300
Qy 352 AAGCTGCTTGGCGATACAGAAATATCATGCTTACATGATGCTAAGGCACACATGAT 411
Db 301 AAGCTGCTTGGCGATACAGAAATATCATGCTTACATGATGCTAAGGCACACATGAT 360
Qy 412 TTCAATGTTTGTGAGAGAGATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTT 471
Db 361 TTCAATGTTTGTGAGAGAGATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTT 420
Qy 472 GCCATGCTGTGTGAGAGAGAGAAATACACTGTTTATTTCTGAAATTCCTGAAATTC 531
Db 421 GCCATGCTGTGTGAGAGAGAGAAATACACTGTTTATTTCTGAAATTCCTGAAATTC 480
Qy 532 AATAGAGCAGCGCTTAAATGCTCTCTTGAAGCCCTCTTTTGGATCTTTTTCAGGCAACA 591
Db 481 AATAGAGCAGCGCTTAAATGCTCTCTTGAAGCCCTCTTTTGGATCTTTTTCAGGCAACA 540
Qy 592 CTGGA CTATGGAATGTATTCGAGAGAGAGAACTATTAAAGAGAGAGAGAGAGAGAGAG 651
Db 541 CTGGA CTATGGAATGTATTCGAGAGAGAGAACTATTAAAGAGAGAGAGAGAGAGAGAG 600
Qy 652 ACAGTCGGAAATTTTACCTTACGATATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 711
Db 601 ACAGTCGGAAATTTTACCTTACGATATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Qy 712 GGATGCGAAATTTTACCTTACGATATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 771
Db 661 GGATGCGAAATTTTACCTTACGATATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Qy 772 AGGCCCAATCTAGTGGAATCTTATACATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 831
Db 721 AGGCCCAATCTAGTGGAATCTTATACATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Qy 832 GCTGATCCAGACTGGATGCTTTTATACATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 891
Db 781 GCTGATCCAGACTGGATGCTTTTATACATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Qy 892 ACCAG 951
Db 841 ACCAG 900
Qy 952 GCCAGATCAGCTGGAGTGGCTACCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1011
Db 901 GCCAGATCAGCTGGAGTGGCTACCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 960
Qy 1012 TATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTA 1071
Db 961 TATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTA 1020
Qy 1072 TATGAG 1131
Db 1021 TATGAG 1080
Qy 1132 ACAAG 1191
Db 1081 ACAAG 1140
Qy 1192 TTTAAGATGTCAGAAATTAATGATGCTGAGAGAGAGAGATCATAGATGTCATAGATAAG 1251
Db 1141 TTTAAGATGTCAGAAATTAATGATGCTGAGAGAGAGAGATCATAGATGTCATAGATAAG 1200
Qy 1252 GAACTAATTCACCTTTTGAGATTTCTATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1311
Db 1201 GAACTAATTCACCTTTTGAGATTTCTATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260

Qy 1312 TGGACTCCTGAGGAAATATGCTTGTGCTCCTCTAGATCGCTCCAGACTCGCCTA 1371
Db 1261 TGGACTCCTGAGGAAATATGCTTGTGCTCCTCTAGATCGCTCCAGACTCGCCTA 1320
Qy 1372 CAGATAGTGTGATCTCTCACTGAATATTTATCCAGTAGAGAGATGATTTGAAAGG 1431
Db 1321 CAGATAGTGTGATCTCTCACTGAATATTTATCCAGTAGAGAGATGATTTGAAAGG 1380
Qy 1432 CAGAGACTCATTTGAGTCTGAGTCTGCTGATCTGTGAGGCACTTAATATCTATGAGAAACA 1491
Db 1381 CAGAGACTCATTTGAGTCTGAGTCTGCTGATCTGTGAGGCACTTAATATCTATGAGAAACA 1440
Qy 1492 ACAGACATCTGTAATATTCATGACATCTTCTCATGTTTTTCCCAAGTCACGAGAG 1551
Db 1441 ACAGACATCTGTAATATTCATGACATCTTCTCATGTTTTTCCCAAGTCACGAGAG 1500
Qy 1552 GAAATGAGTTTATTTTTGCTCTGAAATGCAAAACAGGTTTCCGTCAATTTATACAAAAT 1611
Db 1501 GAAATGAGTTTATTTTTGCTCTGAAATGCAAAACAGGTTTCCGTCAATTTATACAAAAT 1560
Qy 1612 ACATCTATTTTAAAGGAAACAAATATAACGATCCAGTGGTGGCTGCTGCCAAGT 1671
Db 1561 ACATCTA-TTTAAAGGAAACAAATATAACGATCCAGTGGTGGCTGCTGCCAAGT 1619
Qy 1672 GATTTCAAGTCTCTATCAAAGAGAGATAGCAATTTACCAGTGGTGAATTTGGAAGTCTT 1731
Db 1620 GATTTCAAGTCTCTATCAAAGAGAGATAGCAATTTACCAGTGGTGAATTTGGAAGTCTT 1679
Qy 1732 GCGCGGATGGATCAATATCCAAAGTTGATGAAGTCCAGAGGCTGGTATATTTTGAAGG 1791
Db 1680 GCGCGGATGGATCAATATCCAAAGTTGATGAAGTCCAGAGGCTGGTATATTTTGAAGG 1739
Qy 1792 ACCAAAGACTCCCCCTTTAGAGCATCACTGTAGTACGTTACGTTAAATCTCTGAGAG 1851
Db 1740 ACCAAAGACTCCCCCTTTAGAGCATCACTGTAGTACGTTAAATCTCTGAGAG 1799
Qy 1852 GTGCAAGGCTGACTGACCGTGGCTACTCACATCTTCTGTCGATCAGTCAGCAGCTGAC 1911
Db 1800 GTGCAAGGCTGACTGACCGTGGCTACTCACATCTTCTGTCGATCAGTCAGCAGCTGAC 1859
Qy 1912 TTCTTTATAAGTAAATAGTAACTAACAGAGAGATCCACACTGTGTGTCTCTTTACAGCTA 1971
Db 1860 TTCTTTATAAGTAAATAGTAACTAACAGAGAGATCCACACTGTGTGTCTCTTTACAGCTA 1919
Qy 1972 TCAAGTCTGAAGATGACCCAACTTGCRAAACAAAGGAATTTTGGGCCACCATTTTGGAT 2031
Db 1920 TCAAGTCTGAAGATGACCCAACTTGCRAAACAAAGGAATTTTGGGCCACCATTTTGGAT 1979
Qy 2032 TCAGCAGGCTCTTCTTCTGACTATCTCTCTCCAGAAATTTTCTCTTTTGAAGTACTACT 2091
Db 1980 TCAGCAGGCTCTTCTTCTGACTATCTCTCTCCAGAAATTTTCTCTTTTGAAGTACTACT 2039
Qy 2092 GGAATTTACATTTGATGGATGCTCTACAGGCTCATGATCTACAGCTTGAAGAAATAT 2151
Db 2040 GGAATTTACATTTGATGGATGCTCTACAGGCTCATGATCTACAGCTTGAAGAAATAT 2099
Qy 2152 CTTACTGCTCTGTTTATATATGTTGGTCTCTCAGGT 2186
Db 2100 CTTACTGCTCTGTTTATATATGTTGGTCTCTCAGGT 2134

RESULT 23

BD157001

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BD157001 2161 bp DNA linear PAT 17-JAN-2003
Primer for synthesizing full-length cDNA and use thereof.

BD157001.1 GI:27862759

JP 2002191363-A/11844

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.
1 (bases 1 to 2161)
Ota,T., Isegai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 11844 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/11844
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOgai, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ 10, C12P21/02, C12P1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
PT CDS Location/Qualifiers
1..2161
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
Query Match 46.5%; Score 1450; DB 2; Length 2161;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
ORIGIN
832 GCTGATCAGACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTA 891
Db GCTGATCAGACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTA 96
892 ACCAGAGAAGAGAGACTCACTTATGTGCAATGAGCTAGCCAAACATGGAAGAAGAT 951
Db ACCAGAGAAGAGAGAGACTCACTTATGTGCAATGAGCTAGCCAAACATGGAAGAAGAT 156
952 GCCAGATCAGCTGGAGTGGCTACCTTTGTTCTCCAGAGAATTTGATAGATATTTGGC 1011
Db GCCAGATCAGCTGGAGTGGCTACCTTTGTTCTCCAGAGAATTTGATAGATATTTGGC 216
1012 TATTGGTGGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTGAATTTCTA 1071
Db TATTGGTGGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTGAATTTCTA 276
1072 TATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTATCATCCCTATGTTGGAA 1131
Db TATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTATCATCCCTATGTTGGAA 336
1132 ACAAGGGGCGAGATTCAATCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACT 1191
Db ACAAGGGGCGAGATTCAATCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACT 396
1192 TTTTAAGATGTCAGAAATTAATGATTCATGCTGGAAGAGGATCATAGATGTCATAGATAAG 1251
Db TTTTAAGATGTCAGAAATTAATGATTCATGCTGGAAGAGGATCATAGATGTCATAGATAAG 456
1252 GAACTAAATCAACCTTTTGAGATTTCTATTTGAAGAGGTGGAATATATTTGCCAGAGCTGA 1311
Db GAACTAAATCAACCTTTTGAGATTTCTATTTGAAGAGGTGGAATATATTTGCCAGAGCTGA 516
1312 TGGACTCTCGAGGGAATAATGCTTGGTCCATCTACTAGATCGCTCCAGAGCTCGCCTA 1371
Db TGGACTCTCGAGGGAATAATGCTTGGTCCATCTACTAGATCGCTCCAGAGCTCGCCTA 576
1372 CAGATAGTGTGATCTCACTCGAATATTTATCCAGTAGAAGATGATGTTATGGAAGG 1431
Db CAGATAGTGTGATCTCACTCGAATATTTATCCAGTAGAAGATGATGTTATGGAAGG 636

1432 CAGAGACTCATTTGAGTGCAGTGCCTGATTTCTGTGACGCCACTAAATTTATCTATGAAGAAACA 1491
Db CAGAGACTCATTTGAGTGCAGTGCCTGATTTCTGTGACGCCACTAAATTTATCTATGAAGAAACA 696
1492 ACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGTCACGAAGAG 1551
Db ACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGTCACGAAGAG 756
1552 GAAATTGAGTTTATTTTCCCTCTGAATGCAAAAACAGGTTTCCGTCATTATTACAAAATT 1611
Db GAAATTGAGTTTATTTTCCCTCTGAATGCAAAAACAGGTTTCCGTCATTATTACAAAATT 816
1612 ACATCTATTTTAAAGGAAGCAATATTAACCATCCAGTGGTGGCTGCTCCCTCCAGT 1671
Db ACATCTATTTTAAAGGAAGCAATATTAACCATCCAGTGGTGGCTGCTCCCTCCAGT 876
1672 GATTTCAAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTT 1731
Db GATTTCAAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTT 936
1732 GCGCGCATGGATCTTAATATCCAAGTTGATGAAGTCAGAAGCTGGTATATTTTGAAGGC 1791
Db GCGCGCATGGATCTTAATATCCAAGTTGATGAAGTCAGAAGCTGGTATATTTTGAAGGC 996
1792 ACCAAAGACTCCGCTTTAGAGCATCACCTGTAGCTAGTCAGTTACGTAAATCCCTGGAGAG 1851
Db ACCAAAGACTCCGCTTTAGAGCATCACCTGTAGCTAGTCAGTTACGTAAATCCCTGGAGAG 1056
1852 GTCAAGAGCTGATGACCGTGGCTACTCACATTTCTGCTGCACTCAGTCAGCACTGTGAC 1911
Db GTCAAGAGCTGATGACCGTGGCTACTCACATTTCTGCTGCACTCAGTCAGCACTGTGAC 1116
1912 TTCTTTTAAAGTAAGTATAGTAAACCAAGAGATCCACACTGTGTGCTCTTTTACAAGCTA 1971
Db TTCTTTTAAAGTAAGTATAGTAAACCAAGAGATCCACACTGTGTGCTCTTTTACAAGCTA 1176
1972 TCAAGTCTTGAAGATGACCCAACTTTGCAAAAACAAAGGAATTTTGGGCCACCAATTTGGAT 2031
Db TCAAGTCTTGAAGATGACCCAACTTTGCAAAAACAAAGGAATTTTGGGCCACCAATTTGGAT 1236
2032 TCAGCAGTCTCTTCCGCTGATATCTCTCCAGAAATTTCTCTTTTGAAGTACTACT 2091
Db TCAGCAGTCTCTTCCGCTGATATCTCTCCAGAAATTTCTCTTTTGAAGTACTACT 1296
2092 GGAATTACATTTGATGGGATGCTCTACAAGCTCATGATCTACAGCTCGGAAAGAAATAT 2151
Db GGAATTACATTTGATGGGATGCTCTACAAGCTCATGATCTACAGCTCGGAAAGAAATAT 1356
2152 CTTACTGTGCTGTTTCAATATATGTTGGTCTCAGGTGCAAGTGGTGAATTAATCGGTTTAAA 2211
Db CTTACTGTGCTGTTTCAATATATGTTGGTCTCAGGTGCAAGTGGTGAATTAATCGGTTTAAA 1416
2212 GGAAGTCAAGTATTTCCGCTTGAATCCCTAGGCTCTCTAGGTTATGCTGTGTAGTGATA 2271
Db GGAAGTCAAGTATTTCCGCTTGAATCCCTAGGCTCTCTAGGTTATGCTGTGTAGTGATA 1476
2272 GACAAACAGGGGATCTCTGTCAACGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATG 2331
Db GACAAACAGGGGATCTCTGTCAACGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATG 1536
2332 G 2332
1537 G 1537
RESULT 24
AX878058
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
AX878058
Sequence
AX878058
AX878058.1
Homo sapiens (human)
2161 bp
DNA
linear
PAT 17-DEC-2003
12963 from Patent EP1074617.
GI:40032794

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
REFERENCE AUTHORS	Ota, T., Isegai, T., Nishikawa, T., Hayaashi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T. Primers for synthesising full-length cDNA and their use Patent: EP 1074617-A 12963 07-FEB-2001; Research Association for Biotechnology (JP)
FEATURES	Location/Qualifiers 1..2161
source	1..2161
CDS	/organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" 16..1914 /note="unnamed protein product" /codon_start=1 /protein_id="CAE89991.1" /db_xref="GI:40032795" translation="MDPKCFADPDWIAFIHNSDIWISNIVTRERLTYVHNELANN BEDARSAGATFVQDEFDRYGYWPCAEITPDGKILRLIYENDESEVEIHHVT SPMLSTRASFRYPKFTGTANPKVTFMSEIMIDAGRIIDVDEKELQRPPEILPEGV EYAKAGWTEGKYAWSILDRSQRLQVLSPELFPVEDDVMERORLTESVPDSV TPLLIEETDIMINHDIFHPVQSHHEEIEFI FASECKTGFRLHYKITSILKESKY KHLXVSNPSEVRLTRDYSYSHSCCI SQHCDFFISKYQKPNCHVLYKLSPPLE DPTCKTEFWATLDSAGLPDYTPPEIFSPFSTTGTFLYKMLYKPHDLOPGKKYPTV LFYVGGPQVOLVNNRFKGVKFRNLATLASLGVVVVVIDNRGSHRGLKFEKAPFKWV A1AGAPVLMIFDYDTGYTERYMGHPDQEQGYVLGSVAMQAEKFPSPNRLHLHGFL DENVFAHTSILLSFLVRAGKPYDLQIVPQERHSIRVPESGEHVELHLHLQENLGS RIAAALKVI"
ORIGIN	Query Match 46.5%; Score 1450; DB 2; Length 2161; Best Local Similarity 99.9%; Pred. No. 0; Matches 1500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy	832 GCTGATCCAGATCGGATGCTGCTTTATACATAGCAACGATATTTCGATATCTAACATCGTA 891
Db	37 GCTGATCCAGATCGGATGCTGCTTTATACATAGCAACGATATTTCGATATCTAACATCGTA 96
Qy	892 ACCAGAGAAAGAGAGAGATCTCACTTATGTGCACATGAGCTAGCCACACNTGGAGAGAT 951
Db	97 ACCAGAGAAAGAGAGAGATCTCACTTATGTGCACATGAGCTAGCCACACNTGGAGAGAT 156
Qy	952 GCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAGAAATTTGATAGATTTCTGGC 1011
Db	157 GCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAGAAATTTGATAGATTTCTGGC 216
Qy	1012 TATTGTTGTTGTCCTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTAGAAATCTA 1071
Db	217 TATTGTTGTTGTCCTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTAGAAATCTA 276
Qy	1072 TATGAAGAAATGATGAACTGAGGTGGAATTTATGTTATCATGTTACATCCCTATGTTGAA 1131
Db	277 TATGAAGAAATGATGAACTGAGGTGGAATTTATGTTATCATGTTACATCCCTATGTTGAA 336
Qy	1132 ACAAGAGAGGAGATCTATCGTTATCCCTAAAA CAGGTA CAGCAATCTCTAAAGTCACT 1191
Db	337 ACAAGAGAGGAGATCTATCGTTATCCCTAAAA CAGGTA CAGCAATCTCTAAAGTCACT 396
Qy	1192 TTTAGATGTCAGAAATTAATGATGCTGAGGAAGGATCATAGATGTCTAGATAAG 1251
Db	397 TTTAGATGTCAGAAATTAATGATGCTGAGGAAGGATCATAGATGTCTAGATAAG 456
Qy	1252 GAACTAATTCACCTTTTGAGATTCATTTTGAAGGAGTTGAATATATTGCCAGAGCTGGA 1311
Db	457 GAACTAATTCACCTTTTGAGATTCATTTTGAAGGAGTTGAATATATTGCCAGAGCTGGA 516
Qy	1312 TCGACTCTGAGGGGAAATATGCTTGTGCTCCATCTACTAGATCGCTCCAGACTCGCCTA 1371
Db	517 TCGACTCTGAGGGGAAATATGCTTGTGCTCCATCTACTAGATCGCTCCAGACTCGCCTA 576

RESULT 25
AK027826
LOCUS

AK027826 2161 bp mRNA linear PRI 20-JAN-2006

Qy	1372 CAGATAGTGTGATCTCACCCTGAATTTATTTATCCAGTAGAAGATGATGTTATCGAAAGG 1431
Db	577 CAGATAGTGTGATCTCACCCTGAATTTATTTATCCAGTAGAAGATGATGTTATCGAAAGG 636
Qy	1432 CAGAGACTCAATTTGAGTCAAGTCTGATCTCTGAGCGCACCTAATNTATCTATGAAGAAACA 1491
Db	637 CAGAGACTCAATTTGAGTCAAGTCTGATCTCTGAGCGCACCTAATTTATCTATGAAGAAACA 696
Qy	1492 ACAGACATCTGATAAATATCCATGACATCTTTCATGTTTTTCCCAAGTCAAGAGAG 1551
Db	697 ACAGACATCTGATAAATATCCATGACATCTTTCATGTTTTTCCCAAGTCAAGAGAG 756
Qy	1552 GAAATTTGATTTATTTTTCCTCGATTCGAATGCAAAACAGGTTTCGTCATTTATCAAAAT 1611
Db	757 GAAATTTGATTTATTTTTCCTCGATTCGAATGCAAAACAGGTTTCGTCATTTATCAAAAT 816
Qy	1612 ACATCTATTTTAAAGGAAGCAATATAACGATCCAGTGGTGGTGGCTGCTGCTCAAGT 1671
Db	817 ACATCTATTTTAAAGGAAGCAATATAACGATCCAGTGGTGGTGGCTGCTGCTCAAGT 876
Qy	1672 GATTTCAAGTCTCTATCAAAAGAGAGATAGCAATTA CCAAGTGGTGAATGGGAAGTCTT 1731
Db	877 GATTTCAAGTCTCTATCAAAAGAGAGATAGCAATTA CCAAGTGGTGAATGGGAAGTCTT 936
Qy	1732 GCGCGGATGATCTAATATCAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGC 1791
Db	937 GCGCGGATGATCTAATATCAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGC 996
Qy	1792 ACCAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAATCTCGAGAG 1851
Db	997 ACCAAAGACTCCCTTTTAGAGCATCACCTGTAGTCAGTTACGTAATCTCGAGAG 1056
Qy	1852 GTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGTCATCAGTCAGCAGCTGTGAC 1911
Db	1057 GTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGTCATCAGTCAGCAGCTGTGAC 1116
Qy	1912 TTCTTTAATAGTAGTATAAGTAAACAGAAAGATCCACACTGTGTGTCCTTTTACAGCTA 1971
Db	1117 TTCTTTAATAGTAGTATAAGTAAACAGAAAGATCCACACTGTGTGTCCTTTTACAGCTA 1176
Qy	1972 TCAAGTCTGAAAGATGACCCCAACTTCGCAAAACAAAGGAATTTTGGGCCACCATTTGGAT 2031
Db	1177 TCAAGTCTGAAAGATGACCCCAACTTCGCAAAACAAAGGAATTTTGGGCCACCATTTGGAT 1236
Qy	2032 TCAGCAGGTCCTCTTCTGACTATCTCCTCCAGAAATTTTCTCTTTTGAAGTACTACT 2091
Db	1237 TCAGCAGGTCCTCTTCTGACTATCTCCTCCAGAAATTTTCTCTTTTGAAGTACTACT 1296
Qy	2092 GGAATTTACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCTGGAAGAAATAT 2151
Db	1297 GGAATTTACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCTGGAAGAAATAT 1356
Qy	2152 CCTACTGTGCTTTCATATATGTTGGTCTCAGGTGAGTTGGTGAATATCGGTTTAAA 2211
Db	1357 CCTACTGTGCTTTCATATATGTTGGTCTCAGGTGAGTTGGTGAATATCGGTTTAAA 1416
Qy	2212 GGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTTGTTAGTGAATA 2271
Db	1417 GGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTTGTTAGTGAATA 1476
Qy	2272 GACACAGGGGATCTGTGACCGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATG 2331
Db	1477 GACACAGGGGATCTGTGACCGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATG 1536
Qy	2332 G 2332
Db	1537 G 1537

DEFINITION Homo sapiens cDNA FLJ14920 fis, clone PLACE1007416, weakly similar to D1EPR1DYL PEPTIDASE IV (BC 3.4.14.5).

ACCESSION AK027826.1 GI:14042789
VERSION AK027826.1
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Ohtsuka, M., Nishitani, T., Shibahara, T., Tanaka, T., Iehi, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hirao, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Hosoda, M., Horita, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T. O., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Mutsaers, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Mori, Y., Momiya, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, P., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigetani, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isoigai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)

2 Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wagatsuma, M., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahari, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T. NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 2161)

Isogai, T. and Otsuki, T.
Direct Submission
Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: flj-cdna@nifty.com, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

1. 2161
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1007416"

/tissue type="placenta"
/clone_lib="PLACB1"
/note="Cloning vector: pME18SPL3"
16. .1914
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB55395.1"
/db_xref="GI:14042790"
/translation="MDPKLCPADPDWIAFHSDNIWISNIVTREERLTVNHELNM
BEDASAGATFVLQEFDRYSWVWCPKAEITTPSGKILRIDLYEENDESEVILHVT
SPMLSTRADSRFPYKGTANPKVTFKWSEIMIDAEGRIIDVDEKLIQFELFPGV
EYIARAGWTPEGKYAWMSILLDRSQTQLQIVLISPELFIPEVDDVMERQRLIESVPSV
TPLIAYETTDIWINIHDIHFVFPQSHHEEIEFIFASECTGPRHLIKYLSILKESKY
KXSSGLPAPDFKCPKEETAITSGEVLGRHSNIQVDEVRRLVFEFGTKDSSPLE
HHLVYVNVPEVTRTDRSYSHSCISQCHDFFIKSKYNNKPHCVLSYKUGSPED
DPTCKPQWAILDSAGPLPDYTPPEIFSPFESTFTGLTGLYKPHDLQPGKKYPTV
LFIYCGPQVQVNNRFKGVKYPRLNTLASLYVVVVDNRGSHRGLKPFEGAPKXNV
AIAGAPVTIMIFDYGTYERVMGHPDQNEQVYLGVSAMQAEKFPSPNRLLLHGLF
DENVHAFTSILLSLFLVRAGPKYDLYPQERHSIRVPESGEHVELHLHLVQLNLS
RTAALKVI"

ORIGIN

Query Match	46.5%;	Score 1450;	DB 5;	Length 2161;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1500;	Conservative	0;	Mismatches	1;
Indels	0;	Gaps	0;	
QY	832	GCTGATCCGACGATCGATTGCTTTTATACATAGCAACGATATTGGATATCTAAACATCGTA	891	
DB	37	GCTGATCCGACGATCGATTGCTTTTATACATAGCAACGATATTGGATATCTAAACATCGTA	96	
QY	892	ACCAGAGAAGAGGAGACTCATTATGTGCACATAGTAGTACGCCAACATCGAAGAAGAT	951	
DB	97	ACCAGAGAAGAGGAGACTCATTATGTGCACATAGTAGTACGCCAACATCGAAGAAGAT	156	
QY	952	GCCAGATCAGCTGAGTGCCTGCTACTTTGTTTCTTCCAAAGAGAAATTTGATAGATATTTCTGGC	1011	
DB	157	GCCAGATCAGCTGAGTGCCTGCTACTTTGTTTCTTCCAAAGAGAAATTTGATAGATATTTCTGGC	216	
QY	1012	TATTGGTGGTGTCCAAAGCTGAAAACATCCAGTGGTGGTAAATTTCTTAGAATTTCTA	1071	
DB	217	TATTGGTGGTGTCCAAAGCTGAAAACATCCAGTGGTGGTAAATTTCTTAGAATTTCTA	276	
QY	1072	TATGAAGAAATCATGATGATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGAA	1131	
DB	277	TATGAAGAAATCATGATGATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGAA	336	
QY	1132	ACAAGGAGGCGAGATTCATTCCGTTATTCCTTAAACAGGTACAGCAATCTTAAGTCACT	1191	
DB	337	ACAAGGAGGCGAGATTCATTCCGTTATTCCTTAAACAGGTACAGCAATCTTAAGTCACT	396	
QY	1192	TTTAAAGATGTCAGAAATAATGATGATGTCGTGAAGGAGGATCATAGATGTCATAGATAAG	1251	
DB	397	TTTAAAGATGTCAGAAATAATGATGATGTCGTGAAGGAGGATCATAGATGTCATAGATAAG	456	
QY	1252	GAACTAATTCACACCTTTTGAGATTTCTATTTTGAAGGAGTTGAATATATATGCGAGAGCTGA	1311	
DB	457	GAACTAATTCACACCTTTTGAGATTTCTATTTTGAAGGAGTTGAATATATATGCGAGAGCTGA	516	
QY	1312	TGGACTCTCTGAGGGAATAATATGCTTGGTCCATCTCTACTAGATCGTCCAGACTCCGCTTA	1371	
DB	517	TGGACTCTCTGAGGGAATAATATGCTTGGTCCATCTCTACTAGATCGTCCAGACTCCGCTTA	576	
QY	1372	CAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAGG	1431	
DB	577	CAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAGG	636	
QY	1432	CAGAGACTCATTTGAGTGCAGTGCCTGATTTCTGTGACGCCACTAATTTATCTATGAGAAACA	1491	
DB	637	CAGAGACTCATTTGAGTGCAGTGCCTGATTTCTGTGACGCCACTAATTTATCTATGAGAAACA	696	
QY	1492	ACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCCAAGTCCAGAGAG	1551	
DB	697	ACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCCAAGTCCAGAGAG	756	

QY	1552	GAATTTGAGTTATTTTTCCTCTGAAATGCAAAACAGGTTTCGTCAATTTATACAAATTT	1611
Db	757	GAATTTGAGTTATTTTTCCTCTGAAATGCAAAACAGGTTTCGTCAATTTATACAAATTT	816
QY	1612	ACATCTATTTTAAAGGAAGCAATATTAACGATCCAGTGGTGGCTGCTCTCAAGT	1671
Db	817	ACATCTATTTTAAAGGAAGCAATATTAACGATCCAGTGGTGGCTGCTCTCAAGT	876
QY	1672	GATTTCAAGTCTCTATCAAGGAGGATAGCAATTTACAGTGGTGAATGGGAAGTTCTT	1731
Db	877	GATTTCAAGTCTCTATCAAGGAGGATAGCAATTTACAGTGGTGAATGGGAAGTTCTT	936
QY	1732	GGCGGCAATGATCTAATATCCAGTTGATGAAGTCAGAAAGCTGGTATATTTGAAGGC	1791
Db	937	GGCGGCAATGATCTAATATCCAGTTGATGAAGTCAGAAAGCTGGTATATTTGAAGGC	996
QY	1792	ACCAAGACTCCCTTTTAGAGCATCACTGTAGTACAGTACGTTAAATCCCTGGAGAG	1851
Db	997	ACCAAGACTCCCTTTTAGAGCATCACTGTAGTACAGTACGTTAAATCCCTGGAGAG	1056
QY	1852	GTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGCATCAGTCAGCACTGTGAC	1911
Db	1057	GTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGCATCAGTCAGCACTGTGAC	1116
QY	1912	TTCTTTTATAAGTATAGTAAACAGAAAGATCCACATCTGTGTCTCTTTTCAAGCTA	1971
Db	1117	TTCTTTTATAAGTATAGTAAACAGAAAGATCCACATCTGTGTCTCTTTTCAAGCTA	1176
QY	1972	TCAAGTCTTGAAGATGACCCCACTTGCAAAACAAAGGAATTTTGGGCCACCATTTGGAT	2031
Db	1177	TCAAGTCTTGAAGATGACCCCACTTGCAAAACAAAGGAATTTTGGGCCACCATTTGGAT	1236
QY	2032	TCAGCAGGTCCTCTTCTGCTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACT	2091
Db	1237	TCAGCAGGTCCTCTTCTGCTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACT	1296
QY	2092	GGATTTACATTTGATGGGATGCTCTACAGCTCTATGATCTACAGCTTGAAGGAATAT	2151
Db	1297	GGATTTACATTTGATGGGATGCTCTACAGCTCTATGATCTACAGCTTGAAGGAATAT	1356
QY	2152	CCTACTGTCTGTTTATATGTTGCTCTCAGGTGCAAGTGGTGAATATTCGGTTTAA	2211
Db	1357	CCTACTGTCTGTTTATATGTTGCTCTCAGGTGCAAGTGGTGAATATTCGGTTTAA	1416
QY	2212	GGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTTGTAGTGATA	2271
Db	1417	GGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTTGTAGTGATA	1476
QY	2272	GACACAGGGGATCTCTGTCAACGAGGCTTAAATTTGAAGCGCGCTTTAAATATAAAATG	2331
Db	1477	GACACAGGGGATCTCTGTCAACGAGGCTTAAATTTGAAGCGCGCTTTAAATATAAAATG	1536
QY	2332	G	2332
Db	1537	G	1537

RESULT 26
AX405771
LOCUS
DEFINITION
Sequence 186 from Patent WO0222660.
ACCESSION
AX405771.1
VERSION
AX405771.1
KEYWORDS
Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Tang, Y. T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, F.,
Xue, A. J., Yang, Y., Wehrman, T. and Drmanac, R. F.

TITLE	Novel nucleic acids and polypeptides	
JOURNAL	Patent: WO 0222660-A 186 21-MAR-2002;	
FEATURES	HYSEQ, INC. (US)	
source	Location/Qualifiers	
CDS	1..2668 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" 234..2408 /note="unnamed protein product" /codon_start=1 /protein_id="CAD34878.1" /db_xref="GI:21438982" /translations="MAAMETEQIOLGVEIETADCBENIESQDRPKLEPFYVRYWSQ LKKLLASTRIKGLWMAKAPHDPMFKRNDPDPHSDRIYYLAMGENRENTLFTSEI PKTNRAAVLHMSKPLLDLFOQOPLRPNLIVETSCPNIRMDPKLCPADDMTAFTHSN DIWISNIVTREERLIVYHNELANBEDARSAGVATFVLOEBFDYRSGYVWCPKAEIT DSGKILRLIYENDSEVEIHHVTSPLMETRRADSFYKPTGTANPKVTFKMSIMI DAKRIIDVIDKELIQPFELPGEVYIARAGTPEGKYAWSILLDRSQTRLIQIVLIS PELFIPEDDVMERQRLIESVPDSVTPLIYEETDIWINIHDIHFVFPQSHHEEIEF IFASECTQFRHLKYKITSILKSKYKRSGGLPAPDFKCPIKEBIAITSGWEVIEGR HGSNIQVDEVRLLYFEGTKDSPLEHLYVNSVNFGEVTRLDRGYSHSCISQHC FFISKYQXNPHCVSLYKLSPDDPTCKTEFWATILDSAGPLDPTPEIFPFES TTGTLYGMLYKPHDLQPGKKYTVLFYGGPQVAIGAPVTLWTFYDTGVTERVMGH PDNEQSYILGSLVMAQAEKPPSEPNRLLLHLHGFLENVHFAHTSILLSLFVRAGPKPY LQIYPOERHSIRVPESGEHYELHLLHLOENLGSRIAALKVI"	
ORIGIN		
Query Match	44.1%; Score 1377; DB 2; Length 2668;	
Best Local Similarity	99.9%; Pred. No. 0;	
Matches 1427; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
QY	759 GCAACAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCAACATACGATGATCC	818
Db	605 GCAACAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCAACATACGATGATCC	664
QY	819 AAAATTATGCCCGCTGATCCAGACTGGATGCTTTTATACATACACGATATTTGGAT	878
Db	665 AAAATTATGCCCGCTGATCCAGACTGGATGCTTTTATACATACACGATATTTGGAT	724
QY	879 ATCTAAACATCGTAAACAGAGAAGAGAGCTCACTTATGTGCAATAGCATAGCCAA	938
Db	725 ATCTAAACATCGTAAACAGAGAAGAGAGAGCTCACTTATGTGCAATAGCATAGCCAA	784
QY	939 CATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTGTCTCCAAAGAATTTGA	998
Db	785 CATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTGTCTCCAAAGAATTTGA	844
QY	999 TAGATATTCCTGGCTATTTGGTGTGTCCTCAAAAGCTGAAAACAATCCCGATGGTGAAT	1058
Db	845 TAGATATTCCTGGCTATTTGGTGTGTCCTCAAAAGCTGAAAACAATCCCGATGGTGAAT	904
QY	1059 TCTTAGAATCTATATGAAGAAAATGATGAATCTGAGGTGGAATTTATTCATGTTACATC	1118
Db	905 TCTTAGAATCTATATGAAGAAAATGATGAATCTGAGGTGGAATTTATTCATGTTACATC	964
QY	1119 CCTATGTTGAAAACAAGAGGGCAGATTCATTCGGTTATCTCTAAACAGAGTACAGCAA	1178
Db	965 CCTATGTTGAAAACAAGAGGGCAGATTCATTCGGTTATCTCTAAACAGAGTACAGCAA	1024
QY	1179 TCTTAAAGTCACCTTTTAAAGATGTGAGAATAATGATGATGCTGAAGGAAGGATCATAGA	1238
Db	1025 TCTTAAAGTCACCTTTTAAAGATGTGAGAATAATGATGATGCTGAAGGAAGGATCATAGA	1084
QY	1239 TGTCTAGATAAGGAACCTAATTCACCTTTTGAAGATCTATTTGAAGGAGTTGAATATAT	1298
Db	1085 TGTCTAGATAAGGAACCTAATTCACCTTTTGAAGATCTATTTGAAGGAGTTGAATATAT	1144
QY	1299 TGCCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTTGGTCCATCTCTACTAGATCGCTC	1358
Db	1145 TGCCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTTGGTCCATCTCTACTAGATCGCTC	1204
QY	1359 CCAGACTCGCCTACAGATAGTGTGTTGATCTCACCTGAAATTTATTTATCCCAAGTAGAAGATGA	1418

		1. 4309		source		/organism="unknown"		/mol_type="genomic DNA"	
				Query Match		38.2%; Score 1191; DB 2; Length 4309;			
				Best Local Similarity		99.9%; Pred. No. 0;			
				Matches 1371; Conservative		0; Mismatches		0; Indels	
				2; Gaps		1;			
				ORIGIN					
				Qy		1750 ATCCAAAGTTGATGAAGTCAAGAGGCTGGTATATATTTTGAAGGCACCAAGAGCTCCCTTTA		1809	
				Db		1230 ATCCAAAGTTGATGAAGTCAAGAGGCTGGTATATATTTTGAAGGCACCAAGAGCTCCCTTTA		1289	
				Qy		1810 GAGCATCAGTGTACGTAGTACGTAAATCTCGAGAGGTGACAAAGCTGACATGAC		1869	
				Db		1290 GAGCATCAGTGTACGTAGTACGTAAATCTCGAGAGGTGACAAAGCTGACATGAC		1349	
				Qy		1870 CGTGGCTACTCAATCTTTGCTGCTAGTCAAGCTGAGCACTGTGACTTCTTTATTAAGTAAGTAT		1929	
				Db		1350 CGTGGCTACTCAATCTTTGCTGCTAGTCAAGCTGAGCACTGTGACTTCTTTATTAAGTAAGTAT		1409	
				Qy		1930 AGTAACAGAGAAATCCACACTGTGTGTCCCTTTACAGCTATCAAGTCTCAAGATGAC		1989	
				Db		1410 AGTAACAGAGAAATCCACACTGTGTGTCCCTTTACAGCTATCAAGTCTCAAGATGAC		1469	
				Qy		1990 CCAACTGGGAAAACAAAGGAAATTTGGGCCACCATTTTGGATTCAAGCAGGTCTCTTCT		2049	
				Db		1470 CCAACTTTGCAAAAACAAAGGAAATTTGGGCCACCATTTTGGATTCAAGCAGGTCTCTTCT		1529	
				Qy		2050 GACTATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTAATGATGATGGG		2109	
				Db		1530 GACTATACTCTCTCGAATAATTTTCTTTTGAAGTACTACTGGAATTAATGATGATGGG		1589	
				Qy		2110 ATGCTCTACAGCCCTCATGATCTACAGCCCTGGAAAGAAATATCTCTACTGCTGTTTCATA		2169	
				Db		1590 ATGCTCTACAGCCCTCATGATCTACAGCCCTGGAAAGAAATATCTCTACTGCTGTTTCATA		1649	
				Qy		2170 TATGGTGG--TCCTCAGGTGAGTGGTGAATATCGTGTAAAGAGTCAAGTATTTCC		2227	
				Db		1650 TATGGTGGTCTCTCCTCAGGTGAGTGGTGAATATCGTGTAAAGAGTCAAGTATTTCC		1709	
				Qy		2228 GCTTGAATACCTTAGCTCTCTAGTGTATGTTGTTGTTAGTATAGATAGATAGATAGATAGAT		2287	
				Db		1710 GCTTGAATACCTTAGCTCTCTAGTGTATGTTGTTGTTAGTATAGATAGATAGATAGATAGAT		1769	
				Qy		2288 GTCACCGAGGCTTAAATTTGAAGCGCTTTAAATATAAAATGAGTCAAAATAGAAATTTG		2347	
				Db		1770 GTCACCGAGGCTTAAATTTGAAGCGCTTTAAATATAAAATGAGTCAAAATAGAAATTTG		1829	
				Qy		2348 ACGATCAGGTGAAGGATCCCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATC		2407	
				Db		1830 ACGATCAGGTGAAGGATCCCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATC		1889	
				Qy		2408 GTGTGGGATCCACGGCTGCTCCTTAGGAGTACCTCTCCCTGATGGCATTAAATGACGA		2467	
				Db		1890 GTGTGGGATCCACGGCTGCTCCTTAGGAGTACCTCTCCCTGATGGCATTAAATGACGA		1949	
				Qy		2468 GGTTCAGATATCTTCAAGGCTTGTCTATTGCTGGGGCCCAAGTCACTCTGTGATCTTCTATG		2527	
				Db		1950 GGTTCAGATATCTTCAAGGCTTGTCTATTGCTGGGGCCCAAGTCACTCTGTGATCTTCTATG		2009	
				Qy		2528 ATACAGGATACAGGAAAGTTATATGGGTCTACCTTGACAGATGAACAGGGCTATTACT		2587	
				Db		2010 ATACAGGATACAGGAAAGTTATATGGGTCTACCTTGACAGATGAACAGGGCTATTACT		2069	
				Qy		2588 TAGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCT		2647	
				Db		2070 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCT		2129	
				Qy		2648 TACATGTTTCTCGATGAGAAATGTCCATTTTGGCAATACACAGTATATTTACTGATTTTT		2707	
				Db		2130 TACATGTTTCTCGATGAGAAATGTCCATTTTGGCAATACACAGTATATTTACTGATTTTT		2189	

Location/Qualifiers

source		1. .1669	/organism="unknown"		/mol_type="genomic DNA"	
Query Match		32.8%;	Score 1023;	DB 2;	Length 1669;	
Best Local Similarity		100.0%;	Pred. No. 0;			
Matches 1023;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
ORIGIN						
QY	1164	AACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGA	1223			
DB	1	AACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGA	60			
QY	1224	AGGAGGATCATGATGTCATAGATAAGGAACATAATCAACCTTTGAGATTCTATTGGA	1283			
DB	61	AGGAGGATCATGATGTCATAGATAAGGAACATAATCAACCTTTGAGATTCTATTGGA	120			
QY	1284	AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCGAGGAAAATAATGCTGGTCCAT	1343			
DB	121	AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCGAGGAAAATAATGCTGGTCCAT	180			
QY	1344	CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATAT	1403			
DB	181	CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATAT	240			
QY	1404	CCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCAATGAGTCAGTGCCTGATCTGT	1463			
DB	241	CCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCAATGAGTCAGTGCCTGATCTGT	300			
QY	1464	GAGCCCACTAATATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGATCTT	1523			
DB	301	GAGCCCACTAATATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGATCTT	360			
QY	1524	TCATGTTTTTCCCAAGTACGAGAGGAATTTGAGTTATTTTTCCTCTGATGCAAA	1583			
DB	361	TCATGTTTTTCCCAAGTACGAGAGGAATTTGAGTTATTTTTCCTCTGATGCAAA	420			
QY	1584	AACAGGTTTCCGTCATTTATACAAATACATCTATTTTAAAGAAAGCAATATAAGC	1643			
DB	421	AACAGGTTTCCGTCATTTATACAAATACATCTATTTTAAAGAAAGCAATATAAGC	480			
QY	1644	ATCCAGTGTGGCTGCTGCTCCAGTATTTCAAGTGATTTCAAGTGATTTCAAGAGGAGATAGC	1703			
DB	481	ATCCAGTGTGGCTGCTGCTCCAGTATTTCAAGTGATTTCAAGTGATTTCAAGAGGAGATAGC	540			
QY	1704	AATACCAAGTGTGAATGGGAAGTTCTTGGCCGCGATGGATCTTAATATCCAAAGTTGATGA	1763			
DB	541	AATACCAAGTGTGAATGGGAAGTTCTTGGCCGCGATGGATCTTAATATCCAAAGTTGATGA	600			
QY	1764	AGTCAGAGGCTGATATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTA	1823			
DB	601	AGTCAGAGGCTGATATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTA	660			
QY	1824	CGTAGTCAGTTACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA	1883			
DB	661	CGTAGTCAGTTACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA	720			
QY	1884	TTCTTGTGTCATCAGTCAGCACTGTGACTCTTTTAAAGTAAAGTATAGTAACCAAGAA	1943			
DB	721	TTCTTGTGTCATCAGTCAGCACTGTGACTCTTTTAAAGTAAAGTATAGTAACCAAGAA	780			
QY	1944	TCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAAC	2003			
DB	781	TCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAAC	840			
QY	2004	AAAGGAATTTTGGGCCACCACTTTTGGATTCAGCAGGTCTCTCTCTCTGACTACTCTCTCC	2063			
DB	841	AAAGGAATTTTGGGCCACCACTTTTGGATTCAGCAGGTCTCTCTCTCTGACTACTCTCTCC	900			
QY	2064	AGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTCAAGCC	2123			
DB	901	AGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTCAAGCC	960			
ORIGIN						
Query Match		32.8%;	Score 1023;	DB 5;	Length 1669;	
Best Local Similarity		100.0%;	Pred. No. 0;			
Matches 1023;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
ORIGIN						
QY	2124	TCATGATCTACAGCTCGAAGAAATATCTACTGCTGCTTTCATATATGATGCTGCTCA	2183			
DB	961	TCATGATCTACAGCTCGAAGAAATATCTACTGCTGCTTTCATATATGATGCTGCTCA	1020			
QY	2184	GGT 2186				
DB	1021	GGT 1023				
ORIGIN						
RESULT 31						
AF221636						
LOCUS						
DEFINITION	Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA linear PRI 05-NOV-2000					
ACCESSION	AF221636					
VERSION	AF221636.1					
KEYWORDS	GI:11095191					
SOURCES	Homo sapiens (human)					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;					
	Hominidae; Homo.					
REFERENCE	1 (bases 1 to 1669)					
AUTHORS	Abbott,C.A., Yu,D.M., Woollatt,E., Sutherland,G.R., McCaughan,G.W.					
	and Gorrell,M.D.					
TITLE	Cloning, expression and chromosomal localization of a novel human					
JOURNAL	dipeptidyl peptidase (DPP) IV homolog, DPP8					
PUBMED	Eur. J. Biochem. 267 (20), 6140-6150 (2000)					
AUTHORS	2 (bases 1 to 1669)					
TITLE	Direct Submission					
JOURNAL	Submitted (06-JAN-2000) A.W. Morrow Gastroenterology and Liver					
	Centre, Centenary Institute of Cell Biology and Cancer Medicine,					
	Locked Bag No.6, Newtown, Sydney, NSW 2042, Australia					
FEATURES	Location/Qualifiers					
source	1. .1669					
	/organism="Homo sapiens"					
	/mol_type="mRNA"					
	/db_xref="taxon:9606"					
	/chromosome="15"					
	/map="15q22"					
	/clone="T8"					
	/tissue_type="placenta"					
	/clone_lib="lambdaTripleExTM: Clontech"					
gene	-1. 1669					
CDS	/gene="DPP8"					
	/gene="DPP8"					
	/note="serine protease; member of prolyl oligopeptidase S9b					
	family belonging to enzyme clan SC; isoform missing the					
	catalytic serine required for enzyme activity; additional					
	splice variants of this mRNA are found in GenBank					
	Accession Numbers AF221634, AF221635 and AF221637"					
	/codon_start=2					
	/product="dipeptidyl peptidase 8"					
	/protein_id="AAG29768.1"					
	/db_xref="GI:11095192"					
	/translation="TGTPANPKVTFKMSIMIDAEGRIDVIDKELIQPEILFEGVEY					
	IARAGTPGKRYAWSILLDRSQRLQIVLISPELFIPEVDDVWERQRLIESVDSVTP					
	LIIYEETDWINIHDIHFVFPQSHEEIEFIFASECKTGFRHLVYKITSILKESYKR					
	SSGSLPAPDFKCPKEEIAITSGEWEVLRGHSNIQVDEVRLLVYFEGTKDSPLEHH					
	LYVSYVNPGEVRLTRDRGSHSCISQHDFISYKQKPHCVSLYKLSSPDDP					
	CTKTEFWATILDSAGLPDYTPPEIFTSFESTFTLYGMLYKPHDLQPGKYPTVLF					
	TYGQPVAIAGAVTLLWIFDYTYTERYMGHDQNEQGYLGVSAMQAEKFPSEPNRL					
	LLHGFUDENVHFAHTSILLSFLVLRAGKPYDLQIYQPERHSIRVPESGHEVLEHLHY					
	LQENLGSRIAALKVI"					

1164 AACAGGTACAGCAATCTTAAAGTCACCTTTTAAAGATGTCAGAAATAATGATGCTGA 1223
|
|
|
1 AACAGGTACAGCAATCTTAAAGTCACCTTTTAAAGATGTCAGAAATAATGATGCTGA 60
|
|
|
1224 AGGAGGATCATAGATGTCATAGATAAGAACTAAATCAACCTTTTGAGATCTATTGA 1283
|
|
|
61 AGGAGGATCATAGATGTCATAGATAAGAACTAAATCAACCTTTTGAGATCTATTGA 120
|
|
|
1284 AGGAGTGAATATATTGCGAGAGCTGATGAGTCTCTGAGGAGAAATATGCTTGGTCCAT 1343
|
|
|
121 AGGAGTGAATATATTGCGAGAGCTGATGAGTCTCTGAGGAGAAATATGCTTGGTCCAT 180
|
|
|
1344 CCTACTAGATGCTCCAGAGTCTCGCTACAGATAGTGTGATCTCACTGGAATATTAT 1403
|
|
|
181 CCTACTAGATGCTCCAGAGTCTCGCTACAGATAGTGTGATCTCACTGGAATATTAT 240
|
|
|
1404 CCCAGTGAAGATGATGTTATGGAAGGCGAGAGTCAATGAGTCAGTGCCTGATCTGT 1463
|
|
|
241 CCCAGTGAAGATGATGTTATGGAAGGCGAGAGTCAATGAGTCAGTGCCTGATCTGT 300
|
|
|
1464 GAGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCT 1523
|
|
|
301 GAGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCT 360
|
|
|
1524 TCATGTTTTTCCCAAGTCACGAAGAGAAATGAGTTATTTTTTGCCTCTGAATGCAA 1583
|
|
|
361 TCATGTTTTTCCCAAGTCACGAAGAGAAATGAGTTATTTTTTGCCTCTGAATGCAA 420
|
|
|
1584 AACAGTTTCCTGCTATTTATCAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAAG 1643
|
|
|
421 AACAGTTTCCTGCTATTTATCAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAAG 480
|
|
|
1644 ATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTTATCAAAAGAGAGATGAC 1703
|
|
|
481 ATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTTATCAAAAGAGAGATGAC 540
|
|
|
1704 AATTACAGTGGTGAATGGAGTTCTTGGCCGGCATGATCTAATATCAAAAGTTGATGA 1763
|
|
|
541 AATTACAGTGGTGAATGGAGTTCTTGGCCGGCATGATCTAATATCAAAAGTTGATGA 600
|
|
|
1764 AGTCAGAGGCTGTATATTTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACTGTA 1823
|
|
|
601 AGTCAGAGGCTGTATATTTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACTGTA 660
|
|
|
1824 CGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTACA 1883
|
|
|
661 CGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTACA 720
|
|
|
1884 TTCTTGCTGCATGATCAGTCAGCACTGTGACTTCTTTATAAGTAAAGTATAGTAACCAAGAA 1943
|
|
|
721 TTCTTGCTGCATGATCAGTCAGCACTGTGACTTCTTTATAAGTAAAGTATAGTAACCAAGAA 780
|
|
|
1944 TCCACACTGTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAAC 2003
|
|
|
781 TCCACACTGTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAAC 840
|
|
|
2004 AAAGGAATTTTGGGCCACCAATTTTGGATTCAGAGGTCCTCTTCTGACTATCTCTCC 2063
|
|
|
841 AAAGGAATTTTGGGCCACCAATTTTGGATTCAGAGGTCCTCTTCTGACTATCTCTCC 900
|
|
|
2064 AGAAATTTTCTTTTGAAGTACTACTGGAATTTTACATTTGATGGGATGCTCTACAAGCC 2123
|
|
|
901 AGAAATTTTCTTTTGAAGTACTACTGGAATTTTACATTTGATGGGATGCTCTACAAGCC 960
|
|
|
2124 TCATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTGTTTATATATGTTGGTCTCTCA 2183
|
|
|
961 TCATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTGTTTATATATGTTGGTCTCTCA 1020
|
|
|
2184 GGT 2186
|
|
1021 GGT 1023

RESULT 32
CQ721272
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 7206 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
1. .2098
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 28.1%; Score 878; DB 2; Length 2098;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 878; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1747 AATATCCAAGTTGATGAGTCAGAGGCTGCTATATTTTGAAGGCAACAAAGACTCCCT 1806
|
|
|
Db 1217 AATATCCAAGTTGATGAGTCAGAGGCTGCTATATTTTGAAGGCAACAAAGACTCCCT 1276
|
|
|
QY 1807 TTGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGTGAAGGCTGACT 1866
|
|
|
Db 1277 TTGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGTGAAGGCTGACT 1336
|
|
|
QY 1867 GACCGTGGCTACTCACAATTTGCTGCATCAGTCAGTCACTGCTTCTTATTAAGTAAG 1926
|
|
|
Db 1337 GACCGTGGCTACTCACAATTTGCTGCATCAGTCAGTCACTGCTTCTTATTAAGTAAG 1396
|
|
|
QY 1927 TATAGTAAACAGAGAAATCCACACTGTGCTCCCTTTTACAAGCTATCAAGTCTGAAGAT 1986
|
|
|
Db 1397 TATAGTAAACAGAGAAATCCACACTGTGCTCCCTTTTACAAGCTATCAAGTCTGAAGAT 1456
|
|
|
QY 1987 GACCCAACTTGCMAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCCTTT 2046
|
|
|
Db 1457 GACCCAACTTGCMAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCCTTT 1516
|
|
|
QY 2047 CCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGAT 2106
|
|
|
Db 1517 CCTGACTATCTCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGAT 1576
|
|
|
QY 2107 GGGATGCTCTCAAGCCCTCATGATCTACAGCTCGGAAAGAAATATCTACTGTGCTGTT 2166
|
|
|
Db 1577 GGGATGCTCTCAAGCCCTCATGATCTACAGCTCGGAAAGAAATATCTACTGTGCTGTT 1636
|
|
|
QY 2167 ATATATGTTGGTCTCAAGTGCAGTGTGATTAATCGGTTTAAAGAGTCAAGTATTTTC 2226
|
|
|
Db 1637 ATATATGTTGGTCTCAAGTGCAGTGTGATTAATCGGTTTAAAGAGTCAAGTATTTTC 1696
|
|
|
QY 2227 CGCTTGAATACCTTAGCCCTCTCTAGGTTATGTTGGTGTAGTAGAGCAACAGGGATCC 2286
|
|
|
Db 1697 CGCTTGAATACCTTAGCCCTCTCTAGGTTATGTTGGTGTAGTAGAGCAACAGGGATCC 1756
|
|
|
QY 2287 TGTCAACGAGGGCTTAAATTTCAAGCGCTTTAAATATAAATGCGCTCAAAATAGAAAT 2346
|
|
|
Db 1757 TGTCAACGAGGGCTTAAATTTCAAGCGCTTTAAATATAAATGCGCTCAAAATAGAAAT 1816
|
|
|
QY 2347 GAGCATCAGTGGAGAGACTCCCAATATCTAGCTTCTCGATATGATTTCAATGACTTAGAT 2406
|
|
|
Db 1817 GAGCATCAGTGGAGAGACTCCCAATATCTAGCTTCTCGATATGATTTCAATGACTTAGAT 1876

```
Qy 2407 CGTGTGGGATCCAGGCTGGTCCCTATGGAGGATACCTCCCTGATGGCATTAATGCG 2466
Db 1877 CGTGTGGGATCCAGGCTGGTCCCTATGGAGGATACCTCCCTGATGGCATTAATGCG 1936
Qy 2467 AGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCACTCACTCTGTGGATCTTCTAT 2526
Db 1937 AGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCACTCACTCTGTGGATCTTCTAT 1996
Qy 2527 GATACAGGATACACGGAACGTTATATGGGTCACTCCCTGACCAAGATGAAACGGGCTATTAC 2586
Db 1997 GATACAGGATACACGGAACGTTATATGGGTCACTCCCTGACCAAGATGAAACGGGCTATTAC 2056
Qy 2587 TTAGGATCTGTGGCCTGACGACGAGAAAGTTCCCTC 2624
Db 2057 TTAGGATCTGTGGCCTGACGACGAGAAAGTTCCCTC 2094

RESULT 33
AR631282
LOCUS AR631282 1356 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 10 from patent US 6844180.
ACCESSION AR631282
VERSION AR631282.1 GI:59770925
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1356)
AUTHORS Qi,S., Akinsanya,K.O., Riviere,P.J.M. and Junien,J.-L.
TITLES Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 10 18-JAN-2005;
Ferring BV; ;
NLX;

FEATURES
source Location/Qualifiers
1..1356
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 28.1%; Score 877; DB 2; Length 1356;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 927; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAGTGCTAAAGCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTCTTCTTAGTGCGG 60
Db 1 AAGTGCTAAAGCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTCTTCTTAGTGCGG 60
Qy 61 CGTTCCGCCCTGGGTTGTCAACGCGCGCGCGCGGAGGAGCCACTGCAACGAGACCG 120
Db 61 CGTTCCGCCCTGGGTTGTCAACGCGCGCGCGCGGAGGAGCCACTGCAACGAGACCG 120
Qy 121 GAGTGGAGCGCGCGCAGCATGAGCGCGCGCGCGCGCGCGCTCCATAGCGCACGTGCGGACG 180
Db 121 GAGTGGAGCGCGCGCAGCATGAGCGCGCGCGCGCGCGCGCTCCATAGCGCACGTGCGGACG 180
Qy 181 TCCGGCGCGCGCGCGGAGGAGAAATGCAACATGCGACGACCAATGGAACAGAACAG 240
Db 181 TCCGGCGCGCGCGCGGAGGAGAAATGCAACATGCGACGACCAATGGAACAGAACAG 240
Qy 241 CTGGGTGTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGATCGG 300
Db 241 CTGGGTGTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGATCGG 300
Qy 301 CCTAAATTTGAGCCCTTTTATTTGAGCGGTATTTCTGGAGTCAGCTTAAAGAGCTGTT 360
Db 301 CCTAAATTTGAGCCCTTTTATTTGAGCGGTATTTCTGGAGTCAGCTTAAAGAGCTGTT 360
Qy 361 GCCGATACAGAAATATCATGGCTACATGATGGCTAAGCGCACCATGATTTTCATGTTT 420
Db 361 GCCGATACAGAAATATCATGGCTACATGATGGCTAAGCGCACCATGATTTTCATGTTT 420
Qy 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAATCTATTACCTTGCCATGTCT 480
```

```
Db 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAATCTATTACCTTGCCATGTCT 480
Qy 481 GGTGAGAAACAGAGAAATACACTGTTTTATTCTGAAATCCCAAACTATCAATAGAGCA 540
Db 481 GGTGAGAAACAGAGAAATACACTGTTTTATTCTGAAATCCCAAACTATCAATAGAGCA 540
Qy 541 GCAGTCTTAATCCTCTCTTGGAGCCCTCTTTGGATCTTTTCAAGCAACACTGGACTAT 600
Db 541 GCAGTCTTAATCCTCTCTTGGAGCCCTCTTTGGATCTTTTCAAGCAACACTGGACTAT 600
Qy 601 GGAATGTATCTCCGAGAGAGAACTATTAAAGAGAAAGAAACGCAATGGAAACAGTCGGA 660
Db 601 GGAATGTATCTCCGAGAGAGAACTATTAAAGAGAAAGAAACGCAATGGAAACAGTCGGA 660
Qy 661 ATTGCTTCTTACGATATATCACCAGGAAGTGGAACATTTCTGTTTCAAGCCGCTAGTGA 720
Db 661 ATTGCTTCTTACGATATATCACCAGGAAGTGGAACATTTCTGTTTCAAGCCGCTAGTGA 720
Qy 721 ATTATACGCTAAAGATGAGGGCCACAGGATTTACGCAACAACTTTAAGGCCCAAT 780
Db 721 ATTATACGCTAAAGATGAGGGCCACAGGATTTACGCAACAACTTTAAGGCCCAAT 780
Qy 781 CTAGTGGAACTAGTGTCCCAACATACGATCGATCCAAATTTATGCCCGCTGATCCA 840
Db 781 CTAGTGGAACTAGTGTCCCAACATACGATCGATCCAAATTTATGCCCGCTGATCCA 840
Qy 841 GACTGGATTTGCTTTTATACATAGCAACGATATTGGATATCTAAACATCGTAACCCAGAGAA 900
Db 841 GACTGGATTTGCTTTTATACATAGCAACGATATTGGATATCTAAACATCGTAACCCAGAGAA 900
Qy 901 GAAAGGAGACTCACTTATGTGCACAATG 928
Db 901 GAAAGGAGACTCACTTATGTGCACAATG 928

RESULT 34
AX608733
LOCUS AX608733 1356 bp DNA linear PAT 17-FEB-2003
DEFINITION Sequence 10 from Patent W00231134.
ACCESSION AX608733
VERSION AX608733.1 GI:28404302
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE 1
AUTHORS Qi,S., Akinsanya,K.O., Riviere,P.J. and Junien,J.L.
TITLES Novel serine protease genes related to dppiv
JOURNAL Patent: WO 0231134-A 10 18-APR-2002;
Ferring BV (NL)

FEATURES
source Location/Qualifiers
1..1356
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 28.1%; Score 877; DB 2; Length 1356;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 927; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAGTGCTAAAGCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTCTTCTTAGTGCGG 60
Db 1 AAGTGCTAAAGCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTCTTCTTAGTGCGG 60
Qy 61 CGTTCCGCCCTGGGTTGTCAACGCGCGCGCGCGGAGGAGCCACTGCAACGAGACCG 120
Db 61 CGTTCCGCCCTGGGTTGTCAACGCGCGCGCGCGGAGGAGCCACTGCAACGAGACCG 120
Qy 121 GAGTGGAGCGCGCGCAGCATGAGCGCGCGCGCGCGCGCTCCATAGCGCACGTGCGGACG 180
Db 121 GAGTGGAGCGCGCGCAGCATGAGCGCGCGCGCGCGCGCTCCATAGCGCACGTGCGGACG 180
Qy 181 TCCGGCGCGCGCGCGGAGGAGAAATGCAACATGCGACGACCAATGGAACAGAACAG 240
Db 181 TCCGGCGCGCGCGCGGAGGAGAAATGCAACATGCGACGACCAATGGAACAGAACAG 240
Qy 241 CTGGGTGTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGATCGG 300
Db 241 CTGGGTGTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGATCGG 300
Qy 301 CCTAAATTTGAGCCCTTTTATTTGAGCGGTATTTCTGGAGTCAGCTTAAAGAGCTGTT 360
Db 301 CCTAAATTTGAGCCCTTTTATTTGAGCGGTATTTCTGGAGTCAGCTTAAAGAGCTGTT 360
Qy 361 GCCGATACAGAAATATCATGGCTACATGATGGCTAAGCGCACCATGATTTTCATGTTT 420
Db 361 GCCGATACAGAAATATCATGGCTACATGATGGCTAAGCGCACCATGATTTTCATGTTT 420
Qy 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAATCTATTACCTTGCCATGTCT 480
```

1. .1197		/organism="unknown"		/mol_type="genomic DNA"	
source					
ORIGIN					
Query Match	25.3%;	Score 790;	DB 2;	Length 1197;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 790;	Conservative	0;	Mismatches	0;	Indels 0;
Gaps					
2331	GGGTCAAAATAGAAAATGACGATCAGGTGGAAGGACTCCATATATCTAGCTTCTCGATATGA	2390			
Db	GGGTCAAAATAGAAAATGACGATCAGGTGGAAGGACTCCATATATCTAGCTTCTCGATATGA	463			
2391	TTTTCATTGACTTATAGATCGGTGGGCATCCAGCTGGTCTTATGGAGGATACCTCTCCCT	2450			
Db	TTTTCATTGACTTATAGATCGGTGGGCATCCAGCTGGTCTTATGGAGGATACCTCTCCCT	523			
2451	GATGGCATTAAATGACAGGTCAGATATCTTCAAGGTTCTTATGCTGGGGCCCGAGTCAC	2510			
Db	GATGGCATTAAATGACAGGTCAGATATCTTCAAGGTTCTTATGCTGGGGCCCGAGTCAC	583			
2511	TCTGTGATCTTCTATGATACAGGATACAGGAAAGTTATATATGGGTCACTGACAGAA	2570			
Db	TCTGTGATCTTCTATGATACAGGATACAGGAAAGTTATATATGGGTCACTGACAGAA	643			
2571	TGAACAGGCTATTTACTTTAGGATCTGTGCCATGCAAGCAGAAAAGTTCCCTCTGAACC	2630			
Db	TGAACAGGCTATTTACTTTAGGATCTGTGCCATGCAAGCAGAAAAGTTCCCTCTGAACC	703			
2631	AAATCGTTTACTGCTCTTACATGTTTCTGAGTGAATGTCATTTTGACATACACAG	2690			
Db	AAATCGTTTACTGCTCTTACATGTTTCTGAGTGAATGTCATTTTGACATACACAG	763			
2691	TATATTACTGATTTTTTACTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCA	2750			
Db	TATATTACTGATTTTTTACTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCA	823			
2751	GGAGAGACACAGCATAGAGATTCTGAAATCGGAGAACATTTATGAACTGCTTTTGA	2810			
Db	GGAGAGACACAGCATAGAGATTCTGAAATCGGAGAACATTTATGAACTGCTTTTGA	883			
2811	CTACCTTCAAGAAAACCTTGGATCAGCTATTTGCTCTTAAAGTGCATATATTTTGACC	2870			
Db	CTACCTTCAAGAAAACCTTGGATCAGCTATTTGCTCTTAAAGTGCATATATATTTGACC	943			
2871	TGTTGAGACTCTCTGGTATACATCTGCTATTTAAACCAATGAGGAGTTTAACTCAACAG	2930			
Db	TGTTGAGACTCTCTGGTATACATCTGCTATTTAAACCAATGAGGAGTTTAACTCAACAG	1003			
2931	AAAAACAGAAATGATCATCATATTTTGATACCTGCAATGATTAACATCTCTGGAAT	2990			
Db	AAAAACAGAAATGATCATCATATTTTGATACCTGCAATGATTAACATCTCTGGAAT	1063			
2991	AAATGGTGGCCTACGAGGCTTACGGTTGCTGATGATTAATCTTAACCTTAACCCAC	3050			
Db	AAATGGTGGCCTACGAGGCTTACGGTTGCTGATGATTAATCTTAACCTTAACCCAC	1123			
3051	ATGCTCAAAATCAAATGATACATATTTCTGAGAGCCAGCAATACATTAAGAAATTTACTA	3110			
Db	ATGCTCAAAATCAAATGATACATATTTCTGAGAGCCAGCAATACATTAAGAAATTTACTA	1183			
3111	AAAAAAAAAAAA 3120				
Db	AAAAAAAAAAAA 1193				
RESULT 36					
AF221635					
LOCUS	1197 bp	mRNA	linear	PRI 05-NOV-2000	
DEFINITION	Homo sapiens dipeptidyl peptidase 8 (DPP8)	mrna	partial cde,		
ACCESSION	AF221635				
VERSION	AF221635.1	GI:11095189			

```

KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo;
REFERENCE   1 (bases 1 to 1197)
AUTHORS    Abbott C.A., Yu D.M., Woollatt, E., Sutherland, G.R., McCaughan, G.W.
            and Gorrell, M.D.
TITLE      Cloning, expression and chromosomal localization of a novel human
            dipeptidyl peptidase (Dpp) IV homolog, DPP8
JOURNAL    Eur. J. Biochem. 267 (20), 6140-6150 (2000)
PUBMED     11012666
REFERENCE   2 (bases 1 to 1197)
AUTHORS    Abbott C.A., Yu D., McCaughan, G.W. and Gorrell, M.D.
TITLE      Direct Submission
JOURNAL    Submitted (08-JAN-2000) A.W. Morrow Gastroenterology and Liver
            Centre, Centenary Institute of Cell Biology and Cancer Medicine,
            Locked Bag No.6, Newtown, Sydney, NSW 2042, Australia
            Location/Qualifiers
            1..1197
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /chromosome="15"
            /map="15q22"
            /clone="T21"
            /tissue_type="placenta"
            /clone_lib="lambdaTriplexTM: Clontech"
            <1..1197
            /gene="DPP8"
            <1..935
            /gene="DPP8"
            /note="serine protease; member of prolyl oligopeptidase S9b
            family belonging to enzyme clan SC; additional splice
            variants of this mRNA are found in GenBank Accession
            Numbers AF221634, AF221636 and AF221637"
            /codon_start=3
            /product="dipeptidyl peptidase 8"
            /protein_id="AA029767.1"
            /db_xref="GI:11095190"
            /translation="PEGTKDPLEHLVVSYPNGEVBTLDRGYSKCCISQHCDF
            FTSKSNQNPCHVSLYKSPEDDPTCKEFPATILDSAGLPDITPEIFSFEST
            TGTLYGMQLPKDLPQPKYPTVLVYGPQGIEDDVGELQIASKYFIDLLDR
            VGIHWSYGYLKLMLMORSDIPLVALGNAPVWIFIDITGTYRYMHPDNDGGY
            YLGSVAMQAEKPPSPNRLILHGLFDENVHFAHTSILLSFLVRAGKPYDLQIYPQER
            HSIRVPESGEHVELLHLHLQENLGSRIALUKVI"

ORIGIN
Query Match      25.3%; Score 790; DB 5; Length 1197;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 790; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2331 GGGTCAATAGAAATTGAGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATATGA 2390
DB 404 GGGTCAATAGAAATTGAGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATATGA 463

QY 2391 TTTCATTGACTTAGATCTGTGGGATCCACGGTGTCTCTATGGAGGATACCTCTCCCT 2450
DB 464 TTTCATTGACTTAGATCTGTGGGATCCACGGTGTCTCTATGGAGGATACCTCTCCCT 523

QY 2451 GATGGCATTATGACAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCA 2510
DB 524 GATGGCATTATGACAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCA 583

QY 2511 TCTGTGATCTTCTATGATACAGGATACAGGACGTTATATGGGTTCACCTCGACAGACA 2570
DB 584 TCTGTGATCTTCTATGATACAGGATACAGGACGTTATATGGGTTCACCTCGACAGACA 643

QY 2571 TGACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGGAACC 2630
DB 644 TGACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGGAACC 703

RESULT 37
LOCUS      AR651455
DEFINITION Sequence 8 from patent US 6881564.
ACCESSION  AR651455
VERSION     AR651455.1 GI:62795941
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 1083)
AUTHORS     Abbott, C.A. and Gorrell, M.D.
TITLE       Dipeptidyl peptidases
JOURNAL     Patent: US 6881564-A, 8 19-APR-2005;
            The University of Sydney; Sydney;
FEATURES    AUX;
            Location/Qualifiers
            1..1083
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match      25.3%; Score 789; DB 2; Length 1083;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 789; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1755 AGTTGATGACGTAGTACAGAGCGTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCA 1814
DB 295 AGTTGATGACGTAGTACAGAGCGTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCA 354

QY 1815 TCACCTGTACGTAGTACAGTTCAGTAAATCTCTGGAGAGGTCACAAAGGCTGACTGACCGTGG 1874
DB 355 TCACCTGTACGTAGTACAGTTCAGTAAATCTCTGGAGAGGTCACAAAGGCTGACTGACCGTGG 414

QY 1875 CTACTCACATCTTCTGCTGCATCAGTCAAGCTGTGACTTCTTTATAGTAAGTATAGTAA 1934

```

Thu Jun 22 09:04:25 2006

FEATURES
source

Location/Qualifiers
1..1083
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q22"
/clone="scPBWcd3-3-1"
/cell_type="PHA stimulated PBMC"
<1..>1083
/gene="DPP8"
<1..>1083
/gene="DPP8"
/note="serine protease; member of prolyl oligopeptidase S9b family belonging to enzyme clan SC; additional splice variants of this mRNA are found in GenBank Accession Numbers AF221634, AF221635 and AF221636"
/codon_start=2
/product="dipeptidyl peptidase 8"
/protein_id="AAG29769.1"
/translation="FEEDARSAGVATFVLQEFDRYSGYWHCPKAETTPSGKILRLYL
ENDSEVEIIVHTSPMLSTRADSFYQPKGTANPKVTFKMSIEMIDASGRLLIVDEV
RILVFEQTKDSPLSHLYVSYVNPGEVTRLTDGYSHSCCISQHCDFIFSKYSNOK
NPHCVSLYKLSPEDDPTCKEFWATLDSAGPLDPYTPPEIFSFESTFTGLYGLM
YKPHDLQPKKPYTVLFIYGGPOVLNNRFGVKYFRLNTLASLGVVVVVVINDRSC
HRGLKPEGAFKYMGGIIEIDDOVEGLQYLASRYDFIDLDVRVGHWSYGGVLSLMALM
QRSDI PRVAIAGAPVTLMIFDYDTGYT"

gene

CDS

Db	415	CTACTCACAATCTTGTGTCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAA	474
Qy	1935	CCAGAAGAACCCACACTGTGTGTCCTTTACAAAGCTATCAAGTCTCTGAAGATGACCCCAAC	1994
Db	475		534
Qy	1995	TTGCAAAACAAAGGAATTTTGGGCGACCATTTTGGATTGAGTGGATGCTCTTCTCGACTA	2054
Db	535	TTGCAAAACAAAGGAATTTTGGGCGACCATTTTGGATTGAGTGGATGCTCTTCTCGACTA	594
Qy	2055	TACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGGATGCT	2114
Db	595	TACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGGATGCT	654
Qy	2115	CTCAAGGCTCATGATCTACAGCTCGAAGAAATATCTCTACTGTCTTCTATATATGG	2174
Db	655	CTCAAGGCTCATGATCTACAGCTCGAAGAAATATCTCTACTGTCTTCTATATATGG	714
Qy	2175	TGGTCTCTCAGGTGCAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAA	2234
Db	715	TGGTCTCTCAGGTGCAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAA	774
Qy	2235	TACCCTAGGCTCTAGGTATGCTGTAGTGTAGTATAGACAAACAGGGGATCTCTGTCAACG	2294
Db	775	TACCCTAGGCTCTAGGTATGCTGTAGTGTAGTATAGACAAACAGGGGATCTCTGTCAACG	834
Qy	2295	AGGGCTTAAATTTGAAGGCGCTTTAAATATAAATGGTCAATAGAAATGACGATCA	2354
Db	835	AGGGCTTAAATTTGAAGGCGCTTTAAATATAAATGGTCAATAGAAATGACGATCA	894
Qy	2355	GGTGAAGGACTCCATATCTAGTCTTCGATATGATTTCAATTCATCTAGTATGCTGGG	2414
Db	895	GGTGAAGGACTCCATATCTAGTCTTCGATATGATTTCAATTCATCTAGTATGCTGGG	954
Qy	2415	CATCCACGGCTGCTTATGAGGATACCTCTCCCTGATGCTGATTAATGACAGAGTCAGA	2474
Db	955	CATCCACGGCTGCTTATGAGGATACCTCTCCCTGATGCTGATTAATGACAGAGTCAGA	1014
Qy	2475	TATCTTCAGGTTGCTATTTGCTGGGCCCCAGTCATCTCTGTGGATCTTCTATGATACAGG	2534
Db	1015	TATCTTCAGGTTGCTATTTGCTGGGCCCCAGTCATCTCTGTGGATCTTCTATGATACAGG	1074
Qy	2535	ATACACGGA 2543	
Db	1075	ATACACGGA 1083	

RESULT 38	AF221637	1083 bp	mRNA	linear	PRI 05-NOV-2000
LOCUS	Homo sapiens dipeptidyl peptidase 8 (DPP8)	mRNA, partial cds,			
DEFINITION	alternatively spliced.				
ACCESSION	AF221637				
VERSION	AF221637.1	GI:11095193			
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	Abbott,C.A., Yu,D.M., Woollatt,E., Sutherland,G.R., McCaughan,G.W.				
AUTHORS	and Gorrell,M.D.				
TITLE	Cloning, expression and chromosomal localization of a novel human				
JOURNAL	dipeptidyl peptidase (DPP) IV homolog, DPP8				
PUBMED	Eur. J. Biochem. 267 (20), 6140-6150 (2000)				
REFERENCE	11012666				
AUTHORS	2 (bases 1 to 1083)				
TITLE	Abbott,C.A., Yu,D., McCaughan,G.W. and Gorrell,M.D.				
JOURNAL	Direct Submission				
PUBMED	Submitted (06-JAN-2000) A.W. Morrow Gastroenterology and Liver				
AUTHORS	Centre, Centenary Institute of Cell Biology and Cancer Medicine,				
TITLE	Locked Bag No.6, Newtown, Sydney, NSW 2042, Australia				
JOURNAL					

ORIGIN

Query Match	25.3%;	Score 789;	DB 5;	Length 1083;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 789;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1755	AGTTGATGAAGTCAGAGGCTGGTATATTTTCGAGGCCACCAAGACTCCCTTTAGAGCA	1814	
Db	295	AGTTGATGAAGTCAGAGGCTGGTATATTTTCGAGGCCACCAAGACTCCCTTTAGAGCA	354	
QY	1815	TCACCTGTACGTAGTACGTAAATCCTCGAGAGGTGACAAGGCTGACTGACCGTGG	1874	
Db	355	TCACCTGTACGTAGTACGTAAATCCTCGAGAGGTGACAAGGCTGACTGACCGTGG	414	
QY	1875	CTACTCACAATTTCTGCTGCACTCAGTCAGCACTGCTCTTTTATAAGTAAGTATAGTAA	1934	
Db	415	CTACTCACAATTTCTGCTGCACTCAGTCAGCACTGCTCTTTTATAAGTAAGTATAGTAA	474	
QY	1935	CCAGAAGAAATCCACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTCTGAAGTACGCCCAAC	1994	
Db	475	CCAGAAGAAATCCACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTCTGAAGTACGCCCAAC	534	
QY	1995	TTGCAAAACAAAGGAATTTTGGGCGACCATTTTGGATTTCAGCAGGTCTCTTCTCTGACTA	2054	
Db	535	TTGCAAAACAAAGGAATTTTGGGCGACCATTTTGGATTTCAGCAGGTCTCTTCTCTGACTA	594	
QY	2055	TACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCT	2114	
Db	595	TACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCT	654	
QY	2115	CTACAGGCTCATGATCTACAGCTCGAAGAAATATCTCTACTGTCTGTTTATATATAGG	2174	
Db	655	CTACAGGCTCATGATCTACAGCTCGAAGAAATATCTCTACTGTCTGTTTATATATAGG	714	
QY	2175	TGGTCTCTCAGGTGCAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAA	2233	
Db	715	TGGTCTCTCAGGTGCAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAA	774	
QY	2235	TACCTAGGCTCTCTAGGTTTATGCTGTTATAGTATAGACAACAGGGGATCTCTGTCAACG	2299	
Db	775	TACCTAGGCTCTCTAGGTTTATGCTGTTATAGTATAGACAACAGGGGATCTCTGTCAACG	834	
QY	2295	AGGGCTTAAATTTGAAGGCGCTTTTAAATATAAATGGGTCAATAGAAATTCAGCATCA	2354	
Db	835	AGGGCTTAAATTTGAAGGCGCTTTTAAATATAAATGGGTCAATAGAAATTCAGCATCA	894	


```
Qy 2355 GGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATCGTGGG 2414
Db 895 GGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATCGTGGG 954
Qy 2415 CATCCACGGCTGGTCTATGCTATGAGGATACCTCTCCCTGATGGCATTATGCGAGGTCAGA 2474
Db 955 CATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTATGCGAGGTCAGA 1014
Qy 2475 TATCTTCAGGGTTGCTATTTGCTGGGGCCCAAGTCACTCTGCTGATCTTCTATGATACAGG 2534
Db 1015 TATCTTCAGGGTTGCTATTTGCTGGGGCCCAAGTCACTCTGCTGATCTTCTATGATACAGG 1074
Qy 2535 ATACACGGA 2543
Db 1075 ATACACGGA 1083

RESULT 39
AR631286
LOCUS AR631286
DEFINITION Sequence 18 from patent US 6844180.
ACCESSION AR631286
VERSION AR631286.1 GI:59770932
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 832)
  OI.S., Akinsanya,K.O., Riviere,P.J.M. and Junien,J.-L.
  Serine Protease genes related to DPPIV
  Patent: US 6844180-A 18 JAN-2005;
  Ferring BV;;
  NLX;
FEATURES
  source      Location/Qualifiers
            1..832
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
Query Match      24.4%; Score 760; DB 2; Length 832;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCTGCTTCTTAGTGGCG 60
Db 1 AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCTGCTTCTTAGTGGCG 60
Qy 61 CGTTCCGCGCTGGGTTGTACCGCGCCGCGCGAGGAGCCACTGCAACGAGGACCG 120
Db 61 CGTTCCGCGCTGGGTTGTACCGCGCCGCGCGAGGAGCCACTGCAACGAGGACCG 120
Qy 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGCGCCGCTCCATAGCGCAGCTCGGACCG 180
Db 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGCGCCGCTCCATAGCGCAGCTCGGACCG 180
Qy 181 TCCGGCGCGCGCGGGGAGGAAATGCAATGCGCAGCAATGGAACAGAACAG 240
Db 181 TCCGGCGCGCGCGGGGAGGAAATGCAATGCGCAGCAATGGAACAGAACAG 240
Qy 241 CTGGGTGTTGAGATATTGAACTGCGACTGTGAGAGATATTGAATCAGAGTCCG 300
Db 241 CTGGGTGTTGAGATATTGAACTGCGACTGTGAGAGATATTGAATCAGAGTCCG 300
Qy 301 CCTAAATTGGAGCCTTTTATGTTAGCGGTATCTCTGAGTCAGCTTAAAGCTGCTT 360
Db 301 CCTAAATTGGAGCCTTTTATGTTAGCGGTATCTCTGAGTCAGCTTAAAGCTGCTT 360
Qy 361 GCCGATACCGAATAATCATGCTCATGATGGCTAAGGCCACCATGATTTTCATGTTT 420
Db 361 GCCGATACCGAATAATCATGCTCATGATGGCTAAGGCCACCATGATTTTCATGTTT 420
Qy 421 GTGAAGAGGATGATGCAGATGACCTCATTGAGAGAGATCTTACCTTGGCCATGCT 480
```

```
Db 421 GTGAAGAGGATGATGCAGATGGAAGCTCATTGAGAGAGATCTATTACCTTGGCCATGCT 480
Qy 481 GGTGAGAACAGAGAGAAATACACTGTTTTATTCTGAATTCGAAACTATCATATAGAGCA 540
Db 481 GGTGAGAACAGAGAGAAATACACTGTTTTATTCTGAATTCGAAACTATCATATAGAGCA 540
Qy 541 GCAGTCTTAATGCTCTCTTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Qy 601 GGAATGATATTCGAGAGAGAGAACTATTAAAGAGAAAGAAACGCAATGGAACAGTCGGA 660
Db 601 GGAATGATATTCGAGAGAGAGAACTATTAAAGAGAAAGAAACGCAATGGAACAGTCGGA 660
Qy 661 ATTGCTTCTTACGATTATCACAAGGAGTGGAAACATTTCTGTTTCAAGCCGTTAGTGA 720
Db 661 ATTGCTTCTTACGATTATCACAAGGAGTGGAAACATTTCTGTTTCAAGCCGTTAGTGA 720
Qy 721 ATTATTCAGTAAAGATGAGAGGCCACAAGGATTTACGC 760
Db 721 ATTATTCAGTAAAGATGAGAGGCCACAAGGATTTACGC 760

RESULT 40
AX608741
LOCUS AX608741
DEFINITION Sequence 18 from Patent WO0231134.
ACCESSION AX608741
VERSION AX608741.1 GI:28404306
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Hominoidea; Homo.
REFERENCE
  1
  QI.S., Akinsanya,K.O., Riviere,P.J. and Junien,J.L.
  Novel serine protease genes related to DPPIV
  Patent: WO 0231134-A 18 APR-2002;
  Ferring BV (NL)
FEATURES
  source      Location/Qualifiers
            1..832
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match      24.4%; Score 760; DB 2; Length 832;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCTGCTTCTTAGTGGCG 60
Db 1 AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCTGCTTCTTAGTGGCG 60
Qy 61 CGTTCCGCGCTGGGTTGTACCGCGCCGCGCGAGGAGCCACTGCAACGAGGACCG 120
Db 61 CGTTCCGCGCTGGGTTGTACCGCGCCGCGCGAGGAGCCACTGCAACGAGGACCG 120
Qy 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGCGCCGCTCCATAGCGCAGCTCGGACCG 180
Db 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGCGCCGCTCCATAGCGCAGCTCGGACCG 180
Qy 181 TCCGGCGCGCGCGGGGAGGAAATGCAATGCGCAGCAATGGAACAGAACAG 240
Db 181 TCCGGCGCGCGCGGGGAGGAAATGCAATGCGCAGCAATGGAACAGAACAG 240
Qy 241 CTGGGTGTTGAGATATTGAACTGCGACTGTGAGAGATATTGAATCAGAGTCCG 300
Db 241 CTGGGTGTTGAGATATTGAACTGCGACTGTGAGAGATATTGAATCAGAGTCCG 300
Qy 301 CCTAAATTGGAGCCTTTTATGTTAGCGGTATCTCTGAGTCAGCTTAAAGCTGCTT 360
```

Db
203 GATGGCAATTAATGCAGAGGTCAGATATCTTCAGGGTTTGCTATTGCTGGGGCCCCAGTCAC 262

Query Match 22.7%; Score 708; DB 2; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Query Match 22.7%; Score 708; DB 2; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0

2274 CAACAGGGGATCCTGTCCACGAGGCTTAAATTTGAGGCGCTTTAAATATATAAATGGG 2333
Db 1 CAACAGGGGATCCTGTCCACGAGGCTTAAATTTGAGGCGCTTTAAATATATAAATGGG 60
2334 TCAAAATAGAAATTTAGCATCAGGTGGAGGATCCCAATATCTAGCTTCTCGATATGATTT 2393
Db 61 TCAAAATAGAAATTTAGCATCAGGTGGAGGATCCCAATATCTAGCTTCTCGATATGATTT 120
2394 CATTGACATAGATCGTGGGATCCACGGCTGGTCTATGAGGATACCTCTCCCTGAT 2453
Db 121 CATTGACATAGATCGTGGGATCCACGGCTGGTCTATGAGGATACCTCTCCCTGAT 180
2454 GGCATTATGAGAGGATCAGATATCTTCAGGGTGGTATGCTGGGGGCCCAAGTCACTCT 2513
Db 181 GGCATTATGAGAGGATCAGATATCTTCAGGGTGGTATGCTGGGGGCCCAAGTCACTCT 240
2514 GTGGATCTTCTATGATACAGGATACAGGAACTTATATGAGGATACCTCTCCCTGATGA 2573
Db 241 GTGGATCTTCTATGATACAGGATACAGGAACTTATATGAGGATACCTCTCCCTGATGA 300
2574 ACAGGGCTATTAATTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAA 2633
Db 301 ACAGGGCTATTAATTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAA 360
2634 TCGTTTACTCTTACATGTTCTCTGGATGAGATGTCATTTTGACATACAGTAT 2693
Db 361 TCGTTTACTCTTACATGTTCTCTGGATGAGATGTCATTTTGACATACAGTAT 420
2694 ATTAAGTATTTTATGAGGCTGGAAGGATGATGATTTTACAGATCTATCTCCAGGA 2753
Db 421 ATTAAGTATTTTATGAGGCTGGAAGGATGATGATTTTACAGATCTATCTCCAGGA 480
2754 GAGACAGCATTAAGAGTTTCTGAATCGGAGAACTTATGATGATGATTTTACAGATCTATCTCCAGGA 2813
Db 481 GAGACAGCATTAAGAGTTTCTGAATCGGAGAACTTATGATGATGATTTTACAGATCTATCTCCAGGA 540
2814 CCTTCAAGAAACCTTGGATCAGTATGCTGCTCTCAAGTATATTTTACAGTAT 2873
Db 541 CCTTCAAGAAACCTTGGATCAGTATGCTGCTCTCAAGTATATTTTACAGTAT 600
2874 GTAGAACTCTGATATACACTGAGTATTTTACAGAAAGAGGATTTTATCAACAGAA 2933
Db 601 GTAGAACTCTGATATACACTGAGTATTTTACAGAAAGAGGATTTTATCAACAGAA 660
2934 ACAGAAATGATCATCATTTTATGATGATGATGATGATGATGATGATGATGATGAT 2981
Db 661 ACAGAAATGATCATCATTTTATGATGATGATGATGATGATGATGATGATGATGAT 708

RESULT 43
AX524935
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Liou, J.
Regulation of human dipeptidyl peptidase 8
Patent: WO 0206627-A 8 29-AUG-2002;
Bayer Aktiengesellschaft (DE)
Location/Qualifiers
1..753
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN

Query Match 22.4%; Score 700; DB 2; Length 753;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 750; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1609 ATTACATCTATTTTAAAGAAAGCAATATATAACGATCCAGTGGTGGCTGCTCCCA 1668
Db 1 ATTACATCTATTTTAAAGAAAGCAATATATAACGATCCAGTGGTGGCTGCTCCCA 60
1669 AGTGAATTTCAAGTGCCTTCAAGAGAGAGATACCAATACCAAGTGGTGAATGGAGTT 1728
Db 61 AGTGAATTTCAAGTGCCTTCAAGAGAGAGATACCAATACCAAGTGGTGAATGGAGTT 120
1729 CTTGGCCGCGATGGATCTTAATATCCAAAGTTCAGTGAAGTCAAGAGCTGGTATATTGAA 1788
Db 121 CTTGGCCGCGATGGATCTTAATATCCAAAGTTCAGTGAAGTCAAGAGCTGGTATATTGAA 180
1789 GGCACCAAGACTCCCTTTAGAGATCACTGTAGTCACTGATCGTAAATCCTGGA 1848
Db 181 GGCACCAAGACTCCCTTTAGAGATCACTGTAGTCACTGATCGTAAATCCTGGA 240
1849 GAGGTGACAGCTGATGACCGTGGCTACTACATTTCTGTGCTCAGTCACTGCT 1908
Db 241 GAGGTGACAGCTGATGACCGTGGCTACTACATTTCTGTGCTCAGTCACTGCT 300
1909 GACTTTTATAGTAAAGTATAGTAAACCAAGATCCACACTGTGTGCTTTACAAG 1968
Db 301 GACTTTTATAGTAAAGTATAGTAAACCAAGATCCACACTGTGTGCTTTACAAG 360
1969 CTATCAAGTCTGAAAGTACCCCAACTTGCACAAACAAAGAAATTTTGGGCCACCAATTTG 2028
Db 361 CTATCAAGTCTGAAAGTACCCCAACTTGCACAAACAAAGAAATTTTGGGCCACCAATTTG 420
2029 GATTGAGCAGGTCTCTTCTGCTGATCTACTCTCCAGAAATTTCTCTTTGAAAGTACT 2088
Db 421 GATTGAGCAGGTCTCTTCTGCTGATCTACTCTCCAGAAATTTCTCTTTGAAAGTACT 480
2089 ACTGGATTTACATTTGATGGATGCTCTACAAAGCTCTACAGCTCTACAGCTGGAAGAA 2148
Db 481 ACTGGATTTACATTTGATGGATGCTCTACAAAGCTCTACAGCTCTACAGCTGGAAGAA 540
2149 TATCCTACTGCTGTTTATATATATGTTGCTCAGTGGTGGTGAATTAATCGGTTT 2208
Db 541 TATCCTACTGCTGTTTATATATATGTTGCTCAGTGGTGGTGAATTAATCGGTTT 600
2209 AAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTAGTGTATGTTGTTAGTG 2268
Db 601 AAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTAGTGTATGTTGTTAGTG 660
2269 ATAGACACAGGGATCCCTGTCACCGAGGCTTAAATTTGAAGCGCTTTTAAATATAAA 2328
Db 661 ATAGACACAGGGATCCCTGTCACCGAGGCTTAAATTTGAAGCGCTTTTAAATATAAA 720
2329 ATGGGTCAATAGAAATTCAGATCAGGTGG 2359
Db 721 ATGGGTCAATAGAAATTCAGATCAGGTGG 751

RESULT 44
AX524939
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Liou, J.
Regulation of human dipeptidyl peptidase 8

JOURNAL Patent: WO 02066627-A 12 29-AUG-2002;
Bayer Aktiengesellschaft (DE)
Location/Qualifiers
1. .1002
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

FEATURES
source

ORIGIN

Query Match 21.5%; Score 672; DB 2; Length 1002;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2226 CGCTTGAATACCCCTAGCTCTAGCTTATGCTTGTAGTATAGACAAACAGGGATC 2285
Db |||||
QY 15 CGCTTGAATACCCCTAGCTCTAGCTTATGCTTGTAGTATAGACAAACAGGGATC 74
Db |||||
QY 2286 CTGTCCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAATAGAAAT 2345
Db |||||
QY 75 CTGTCCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAATAGAAAT 134
QY 2346 TGACGATCAGGTGAAGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTGACTTGA 2405
Db |||||
QY 135 TGACGATCAGGTGAAGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTGACTTGA 194
QY 2406 TCGTGTGGGCATCCACGCGTGGTCTATGAGGATACCTCTCCCTGATGGCATTAATGCA 2465
Db |||||
QY 195 TCGTGTGGGCATCCACGCGTGGTCTATGAGGATACCTCTCCCTGATGGCATTAATGCA 254
QY 2466 GAGTTCAGATATCTTACGGTGTCTATGCTGGGGCCCAAGTCACTCTGTGGATCTTCTA 2525
Db |||||
QY 255 GAGTTCAGATATCTTACGGTGTCTATGCTGGGGCCCAAGTCACTCTGTGGATCTTCTA 314
QY 2526 TGATACAGGATACACGGAACGTTATATGCGGTCACTCCCTGACCAAGTGAACAGGGCTATT 2585
Db |||||
QY 315 TGATACAGGATACACGGAACGTTATATGCGGTCACTCCCTGACCAAGTGAACAGGGCTATT 374
QY 2586 CTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTGCT 2645
Db |||||
QY 375 CTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTGCT 434
QY 2646 CTTACATGGTTTCTCGATGAGATGTCATTTTGCACATACCAAGTATATTACTGATTT 2705
Db |||||
QY 435 CTTACATGGTTTCTCGATGAGATGTCATTTTGCACATACCAAGTATATTACTGATTT 494
QY 2706 TTTAGTGGGCTGGAAGCCATGATGTTTACAGATCTATCTCCAGAGAGACACAGCAT 2765
Db |||||
QY 495 TTTAGTGGGCTGGAAGCCATGATGTTTACAGATCTATCTCCAGAGAGACACAGCAT 554
QY 2766 AAGAGTTCCTGAATCGGGAGAACATTAATGAATGCACTGATCTTTTGCACATACCTTCAAGAAA 2825
Db |||||
QY 555 AAGAGTTCCTGAATCGGGAGAACATTAATGAATGCACTGATCTTTTGCACATACCTTCAAGAAA 614
QY 2826 CTTTGGATCAGTATGCTGCTCTAAAGTGATATAATTTTGAACCTGTGAGACTCTCT 2885
Db |||||
QY 615 CTTTGGATCAGTATGCTGCTCTAAAGTGATATAATTTTGAACCTGTGAGACTCTCT 674
QY 2886 GGTATACACTGG 2897
Db |||||
QY 675 GGTATACACTGG 686

RESULT 45
BD149852
LOCUS BD149852 631 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD149852
VERSION BD149852.1 GI:27855610
KEYWORDS JP 2002191363-A/4695.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homnidae; Homo.
1 (bases 1 to 631)
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 4695 09-JUL-2002,
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/4695
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ 10,
C12N5/00, C12P21/02, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source 1. .631
/organism="Homo sapiens (human)"
FT Location/Qualifiers
1. .631
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source

ORIGIN

Query Match 16.1%; Score 501; DB 2; Length 631;
Best Local Similarity 99.8%; Pred. No. 1e-282;
Matches 551; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 832 GCTGATCCAGACTGGTTCCTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTA 891
Db |||||
QY 37 GCTGATCCAGACTGGTTCCTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTA 96
Db |||||
QY 892 ACCAGAGAGAAAGAGAGACTCACTTATGTGCACATAGCTAGTAGCAACATCGAAGAGAT 951
Db |||||
QY 97 ACCAGAGAGAAAGAGAGACTCACTTATGTGCACATAGCTAGTAGCAACATCGAAGAGAT 156
QY 952 GCAGATACAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATCTGGC 1011
Db |||||
QY 157 GCAGATACAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATCTGGC 216
QY 1012 TATTGTGTGTCTCCAAAGCTGAAACAACTCCAGTGTGTGTAATAATTTCTAGAAATCTTA 1071
Db |||||
QY 217 TATTGTGTGTCTCCAAAGCTGAAACAACTCCAGTGTGTGTAATAATTTCTAGAAATCTTA 276
QY 1072 TATGAAGAAAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCTTATGTGGAA 1131
Db |||||
QY 277 TATGAAGAAAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCTTATGTGGAA 336
QY 1132 ACAAGAGAGGCGAGATTCATTCGTTATCTTAAACAGGTACAGCAAAATCCTTAAAGTCACT 1191
Db |||||
QY 337 ACAAGAGAGGCGAGATTCATTCGTTATCTTAAACAGGTACAGCAAAATCCTTAAAGTCACT 396
QY 1192 TTTAAGATGTCAGAAAAATATGATTTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAG 1251
Db |||||
QY 397 TTTAAGATGTCAGAAAAATATGATTTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAG 456
QY 1252 GAATTAATTCACCTTTTGGAGTTCTATTTGAGGAGTTGAATATATTTGCCAGAGCTGGA 1311
Db |||||
QY 457 GAATTAATTCACCTTTTGGAGTTCTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGA 516
QY 1312 TGGACTCTCGAGGAAAAATATGCTTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCTA 1371
Db |||||
QY 517 TGGACTCTCGAGGAAAAATATGCTTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCTA 576
QY 1372 CAGATAGTGTG 1383
Db |||||
QY 577 CAGATAGTGTG 588

RESULT 46
 AX869790
 LOCUS
 DEFINITION
 AX869790 Sequence 4695 from Patent EP1074617. 631 bp DNA linear PAT 17-DEC-2003
 ACCESSION
 AX869790
 VERSION
 AX869790.1 GI:40024653
 KEYWORDS
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 REFERENCE
 1 Ota, T., Iwagaki, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
 Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K., and Otsuki, T.
 Primers for synthesizing full-length cDNA and their use
 Patent: EP 1074617-A 4695 07-FEB-2001;
 JOURNAL
 Research Association for Biotechnology (JP)
 FEATURES
 Location/Qualifiers
 1..631
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 16.1%; Score 501; DB 2; Length 631;
 Best Local Similarity 99.8%; Pred. No. 1e-282;
 Matches 551; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 832 GCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATTGATATCTAACATCGTA 891
 Db 37 GCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATTGATATCTAACATCGTA 96
 QY 892 ACCAGAGAGAGAGAGAGACTCACTATGTGCACATAGCTAGCCACATGGAGAGAT 951
 Db 97 ACCAGAGAGAGAGAGAGACTCACTATGTGCACATAGCTAGCCACATGGAGAGAT 156
 QY 952 GCCAGATCAGCTGGAGTCCCTACCTTTGTTCTCCAGAGAGAAATTTGATATTTGGC 1011
 Db 157 GCCAGATCAGCTGGAGTCCCTACCTTTGTTCTCCAGAGAGAAATTTGATATTTGGC 216
 QY 1012 TATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGTGTAATTTCTAGAAATCTTA 1071
 Db 217 TATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGTGTAATTTCTAGAAATCTTA 276
 QY 1072 TATGAAGAAATGATGAATCTGAGGTGGAATTTATGTTACATCCCTATGTTGGAA 1131
 Db 277 TATGAAGAAATGATGAATCTGAGGTGGAATTTATGTTACATCCCTATGTTGGAA 336
 QY 1132 ACAAGAGGCGAGATTCATTCGTTATCTCTAAACAGGTACAGCAATCTTAAAGTCACT 1191
 Db 337 ACAAGAGGCGAGATTCATTCGTTATCTCTAAACAGGTACAGCAATCTTAAAGTCACT 396
 QY 1192 TTTAGATGTCAGAAATATGATTTGATGCTGAGGAGGATCATAGATGTCATAGATAG 1251
 Db 397 TTTAGATGTCAGAAATATGATTTGATGCTGAGGAGGATCATAGATGTCATAGATAG 456
 QY 1252 GAACCTAATCAACCTTTTGAGATTCATTTGAGAGGATTTGAATATATTTGCCAGAGCTGA 1311
 Db 457 GAACCTAATCAACCTTTTGAGATTCATTTGAGAGGATTTGAATATATTTGCCAGAGCTGA 516
 QY 1312 TGGACTCTGAGGAGAAATATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCCTA 1371
 Db 517 TGGACTCTGAGGAGAAATATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCCTA 576
 QY 1372 CAGATAGTGTG 1383
 Db 577 CAGATAGTGTG 588

RESULT 47

AR631285
 LOCUS
 DEFINITION
 AR631285 Sequence 16 from patent US 6844180. 620 bp DNA linear PAT 14-FEB-2005
 ACCESSION
 AR631285
 VERSION
 AR631285.1 GI:59770931
 KEYWORDS
 Unknown.
 SOURCE
 Unknown.
 ORGANISM
 Unclassified.
 REFERENCE
 1 (bases 1 to 620)
 AUTHORS
 O. S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.
 TITLE
 Serine protease genes related to DPPIV
 JOURNAL
 Patent: US 6844180-A 16 18-JAN-2005;
 Ferring BV,;
 NLX,
 FEATURES
 Location/Qualifiers
 1..620
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Query Match 15.1%; Score 472; DB 2; Length 620;
 Best Local Similarity 100.0%; Pred. No. 1.4e-265;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGTCTTAAGCCTCCGAGCCCAAGCCGCTGCTACTGCGCGGCTGCTTCTTAGTGCGG 60
 Db 1 AAGTCTTAAGCCTCCGAGCCCAAGCCGCTGCTACTGCGCGGCTGCTTCTTAGTGCGG 60
 QY 61 CGTTGCGCGGCTGGGTTGTCAACCGCGCGCCGCGAGGAAGCCACTGCAACGAGACG 120
 Db 61 CGTTGCGCGGCTGGGTTGTCAACCGCGCGCCGCGAGGAAGCCACTGCAACGAGACG 120
 QY 121 GAGTGGAGGCGCGCGAGCATGAAGCGCGCGCGCGAGGAAGCCACTGCAACGAGACG 180
 Db 121 GAGTGGAGGCGCGCGAGCATGAAGCGCGCGCGCGAGGAAGCCACTGCAACGAGACG 180
 QY 181 TCCGGGCGCGCGCGCGAGGAAGGAATGCAACATGCGAGCAGCAATGGAACGAGACG 240
 Db 181 TCCGGGCGCGCGCGCGAGGAAGGAATGCAACATGCGAGCAGCAATGGAACGAGACG 240
 QY 241 CTGGGTGTTGAGATATTTGAAACTGCGACTGCTGAGAGGAGATTTGATCAGAGATCG 300
 Db 241 CTGGGTGTTGAGATATTTGAAACTGCGACTGCTGAGAGGAGATTTGATCAGAGATCG 300
 QY 301 CCTAAATTGAGGCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGAGTCTTT 360
 Db 301 CCTAAATTGAGGCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGAGTCTTT 360
 QY 361 GCCGATACCAAGAAATATCATGCTACATGCTGCTAAGGCCACCAATGATTTTATGTT 420
 Db 361 GCCGATACCAAGAAATATCATGCTACATGCTGCTAAGGCCACCAATGATTTTATGTT 420
 QY 421 GTGAAGAGGAATGATCCAGATGAGCTCATTCAGACAGAAATCTATTACCTTG 472
 Db 421 GTGAAGAGGAATGATCCAGATGAGCTCATTCAGACAGAAATCTATTACCTTG 472

RESULT 48
 AX608739
 LOCUS
 DEFINITION
 AX608739 Sequence 16 from Patent WO0231134. 620 bp DNA linear PAT 17-FEB-2003
 ACCESSION
 AX608739.1 GI:28404305
 KEYWORDS
 Homo sapiens (human)
 SOURCE
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 REFERENCE
 1 Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.
 Novel serine protease genes related to dppiv

JOURNAL	Patent: WO 0231134-A 16 18-APR-2002;
FEATURES	Ferring BV (NL) Location/Qualifiers 1. .620 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"
ORIGIN	Query Match 15.1%; Score 472; DB 2; Length 620; Best Local Similarity 100.0%; Pred. No. 1.4e-265; Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 AAGTCTAAAGCTCCGAGGCGCAAGCGCTGTACTGCGCGCTGCTTCTTAGTGCAG 60 Db 1 AAGTCTAAAGCTCCGAGGCGCAAGCGCGCTGTACTGCGCGCTGCTTCTTAGTGCAG 60
QY	61 GCTTCCGCGCTGCTGGTGTGTACCCGCGCCGCGCGAGGAGCACTGCAACGAGACCG 120 Db 61 GCTTCCGCGCTGCTGGTGTGTACCCGCGCCGCGCGAGGAGCACTGCAACGAGACCG 120
QY	121 GAGTGGAGCGCGCGCAGCATGAAGGGCGCGCGCGCTCCATAGCGCACGTCGGGACGG 180 Db 121 GAGTGGAGCGCGCGCAGCATGAAGGGCGCGCGCGCTCCATAGCGCACGTCGGGACGG 180
QY	181 TCCGGCGCGCGCGCGGGAAGGAATGCAACATGCGCAGCAATGGAACAGAAACAGAACAG 240 Db 181 TCCGGCGCGCGCGCGGGAAGGAATGCAACATGCGCAGCAATGGAACAGAAACAGAACAG 240
QY	241 CTGGGTGTGAGATATTTGAAACTCGGACCTGTGAGGAGGAATATTGAAATCAGAGATCGG 300 Db 241 CTGGGTGTGAGATATTTGAAACTCGGACCTGTGAGGAGGAATATTGAAATCAGAGATCGG 300
QY	301 CCTAAATTCGAGCGCTTTTATGTTGAGCGGTATTCCTCGAGTCAGCTTAAAGAGCTGCT 360 Db 301 CCTAAATTCGAGCGCTTTTATGTTGAGCGGTATTCCTCGAGTCAGCTTAAAGAGCTGCT 360
QY	361 GCGGATACAGAAATATCATGCTACATGATGCTAGGACCCACACATGATTTTCATGTTT 420 Db 361 GCGGATACAGAAATATCATGCTACATGATGCTAGGACCCACACATGATTTTCATGTTT 420
QY	421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTG 472 Db 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTG 472
RESULT 49	
AC105129	198295 bp DNA linear PRI 26-APR-2002
LOCUS	Homo sapiens chromosome 15, clone RP11-349G13, complete sequence.
DEFINITION	AC105129
ACCESSION	AC105129.4 GI:20330981
VERSION	HTG.
KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 198295)
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Hagos,B., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczkzy,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Risse,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE	Submitted (26-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
JOURNAL	Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS	On Apr 26, 2002 this sequence version replaced gi:20148056. All repeats were identified using RepeatMasker: Smit, A.P.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Genome Center Center: Whitehead Institute/ MIT Center for Genome Research

MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Risse,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (23-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 198295)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczkzy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Risse,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (14-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 198295)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczkzy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Risse,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (26-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 26, 2002 this sequence version replaced gi:20148056.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22971
Center clone name: 349 G 13

Direct Submission
Submitted (11-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 11, 2002 this sequence, version replaced gi:14994158.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center -----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: t2618
Center clone name: 346_A_8

Only the last 147.8 kb of this clone are being submitted.
The remainder overlaps accession number AC105129 [WICGR project
L22971].

FEATURES	Location/Qualifiers
source	1. .14760
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="15"
	/map="15"
	/clone="RP11-346A8"
repeat_region	572. .672
	/rpt_family="MIR"
repeat_region	697. .1138
	/rpt_family="L1MB6"
repeat_region	1143. .1340
	/rpt_family="AluJo"
repeat_region	1357. .1665
	/rpt_family="AluSg"
repeat_region	1666. .1699
	/rpt_family="L2"
repeat_region	1771. .2026
	/rpt_family="AluJo"
repeat_region	2049. .2069
	/rpt_family="(TC)n"
repeat_region	2072. .2096
	/rpt_family="(CA)n"
repeat_region	2154. .2185
	/rpt_family="(TC)n"
repeat_region	complement(2458. .2533)
	/rpt_family="MIR3"
repeat_region	2943. .3081
	/rpt_family="MIR3"
repeat_region	complement(3473. .3526)
	/rpt_family="L2"
repeat_region	3584. .3758
	/rpt_family="MIR"
repeat_region	complement(3767. .4021)
	/rpt_family="AluSx"
repeat_region	complement(4075. .4115)
	/rpt_family="AluSx"
repeat_region	complement(4737. .4774)
	/rpt_family="MIR"
repeat_region	5062. .5168
	/rpt_family="MIR"
repeat_region	complement(5262. .5563)
	/rpt_family="AluY"
repeat_region	complement(6726. .7022)
	/rpt_family="AluSx"
repeat_region	7049. .7166
	/rpt_family="L2"
repeat_region	7215. .7332
	/rpt_family="MIR"
repeat_region	complement(7387. .7530)
	/rpt_family="L3"

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 147760)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 15, clone RP11-346A8
Unpublished

2 (bases 1 to 147760)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Cattle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission
Submitted (15-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 147760)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
Lamazares,R., Landers,T., Lehoczyk,P., Major,J., Marquis,N.,
Liu,G., MacLean,C., Macdonald,P., McEwan,P., McKernan,K., Meldrim,J.,
Matthews,C., McCarthy,M., Norman,C.H., Murphy,T., Naylor,J., Nguyen,C.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Taylor,J., Triggilo,J.,
Travers,M., Topham,K., Travers,M., Travis,N., Triggilo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (10-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 147760)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,K., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Triggilo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (10-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
JOURNAL
AUTHORS

```
repeat_region      8050..8168
                    /rpt_family="MIR3"
repeat_region      complement(8169..8480)
                    /rpt_family="AluJo"
repeat_region      8481..8537
                    /rpt_family="MIR3"
repeat_region      complement(9062..9373)
                    /rpt_family="AluSx"
repeat_region      9383..9601
                    /rpt_family="L2"
repeat_region      9755..10149
                    /rpt_family="L2"
repeat_region      11647..11688
                    /rpt_family="L2"
repeat_region      11960..11991
                    /rpt_family="CA)n"
repeat_region      12008..12028
                    /rpt_family="QC-rich"
repeat_region      12084..12161
                    /rpt_family="(CCG)n"
repeat_region      complement(12722..12877)
                    /rpt_family="G-rich"
repeat_region      13023..13086
                    /rpt_family="MIR3"
repeat_region      13212..13282
                    /rpt_family="L3"
repeat_region      13283..13537
                    /rpt_family="L2"
repeat_region      13538..13667
                    /rpt_family="L2"
repeat_region      13668..14227
                    /rpt_family="FLAM_C"
repeat_region      complement(14228..14513)
                    /rpt_family="AluSx"
repeat_region      14514..14525
                    /rpt_family="L2"
repeat_region      complement(14526..14835)
                    /rpt_family="AluSx"
repeat_region      14836..14999
                    /rpt_family="L2"
repeat_region      complement(15356..15637)
                    /rpt_family="L2"
repeat_region      15859..15991
                    /rpt_family="MIR3"
repeat_region      16064..16363
                    /rpt_family="AluSq"

Query Match      10.9%; Score 339; DB 5; Length 147760;
Best Local Similarity 100.0%; Pred. No. 4.8e-187;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2782 GGAGAACATTATGACTGCATCTTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTAT 2841
Db      36082 GGAGAACATTATGACTGCATCTTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTAT 36023

QY      2842 GCTGCTCTAAAGTGATATATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTAT 2901
Db      36022 GCTGCTCTAAAGTGATATATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTAT 35963

QY      2902 TTACCAATATGAGAGGTTTAAATCAACAGAAAACACAGAAATTGATCATCATATTTTGATA 2961
Db      35962 TTACCAATATGAGAGGTTTAAATCAACAGAAAACACAGAAATTGATCATCATATTTTGATA 35903

QY      2962 CTGCGCATGTAACATCTACTCTGTAATAAATATGTTGCGCATGACGGGTTCTACGGTTT 3021
Db      35902 CTGCGCATGTAACATCTACTCTGTAATAAATATGTTGCGCATGACGGGTTCTACGGTTT 35843

QY      3022 GTGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGA 3081
Db      35842 GTGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGA 35783

QY      3082 GAGACCCAGCAATACCAATAAGAATTACTAAAAA 3120
                    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      35782 GAGACCCAGCAATACCAATAAGAATTACTAAAAA 35744

RESULT 51
AR263993
LOCUS      AR263993
DEFINITION Sequence 171 from patent US 6331427.
ACCESSION AR263993
VERSION   AR263993.1
KEYWORDS  AR263993.1 GI:28075997
SOURCE    Unknown.
ORGANISM  Unknown.
REFERENCE 1 (bases 1 to 823)
AUTHORS  Robinson,K.E.
TITLE    Protease homologs
JOURNAL  Patent: US 6331427-A 171 18-DEC-2001;
        Millennium Pharmaceuticals, Inc.; Cambridge, MA
FEATURES  Location/Qualifiers
           source          1..823
                        /mol_type="genomic DNA"
ORIGIN

Query Match      9.9%; Score 309; DB 2; Length 823;
Best Local Similarity 99.7%; Pred. No. 2.2e-169;
Matches 359; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2366 TCCAAATATCTAGCTTCTCGATATGATTTCAITGACTTAGATCGTGTGGCATCCACGGCT 2425
Db      35 TCCAAATATCTAGCTTCTCGATATGATTTCAITGACTTAGATCGTGTGGCATCCACGGCT 94

QY      2426 GGTCTATGAGGATACCTCTCCCTGATGGCAATTAATGAGAGGTCAGATATCTTCAGGG 2485
Db      95 GGTCTATGAGGATACCTCTCCCTGATGGCAATTAATGAGAGGTCAGATATCTTCAGGG 154

QY      2486 TTGCTATTGCTGGGGCCCGAGTCACCTCTGTGATCTTCTATGATACAGATACAGGAC 2545
Db      155 TTGCTATTGCTGGGTCCCGAGTCACCTCTGTGATCTTCTATGATACAGATACAGGAC 214

QY      2546 GTTATATGGGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCATGC 2605
Db      215 GTTATATGGGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCATGC 274

QY      2606 AAGCAGAAAAAGTTCCCTCTGTAACCAAAATCGTTTACTGCTCTTACATGTTTCTGGATG 2665
Db      275 AAGCAGAAAAAGTTCCCTCTGTAACCAAAATCGTTTACTGCTCTTACATGTTTCTGGATG 334

QY      2666 AGAATGTCCATTGTCACATACCACTATATTAAGTGGTGTGAGTTTGTAGGGCTGGAAGC 2725
Db      335 AGAATGTCCATTGTCACATACCACTATATTAAGTGGTGTGAGTTTGTAGGGCTGGAAGC 394

RESULT 52
BD154822/c
LOCUS      BD154822
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD154822
VERSION   BD154822.1
KEYWORDS  JP 2002191363-A/9665.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
        Bkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
        Homnidae; Homo.
        1 (bases 1 to 587)
REFERENCE 1
AUTHORS  Ota,i., Isegai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
        Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE    Primer for synthesizing full-length cDNA and use thereof
JOURNAL  Patent: JP 2002191363-A 9665 09-JUL-2002;
        HELIX RESEARCH INSTITUTE
COMMENT   OS Homo sapiens (human)
        PN 2002191363-A/9665
```

```
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC
C12N15/09, C07K14/47, C07K16/19, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10,
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source 1..587
FT /organism="Homo sapiens (human)".

FEATURES
source
Location/Qualifiers
1..587
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 6.6%; Score 205; DB 2; Length 587;
Best Local Similarity 99.2%; Pred. No. 5.1e-108;
Matches 355; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2655 TTTCTCGATGAGAAATGTCATTTTGCACATACCATATATTACTGAGTTTTTTAGTGAG 2714
DB 455 TTTCTCGATGAGAAATGTCATTTTGCACATACCATATATTACTGAGTTTTTTAGTGAG 396
QY 2715 GGCTGGAAGCCATATGATTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCC 2774
DB 395 GGCTGGAAGCCATATGATTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCC 336
QY 2775 TGAATCGGAGAACATTTATGAATCTGATCTTTTGCATCTTCAAGAAAACCTTGGATC 2834
DB 335 TGAATCGGAGAACATTTATGAATCTGATCTTTTGCATCTTCAAGAAAACCTTGGATC 276
QY 2835 ACGTATTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATACAC 2894
DB 275 ACGTATTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTTNTGGTATACAC 216
QY 2895 TGGCTATTTAACCAATGAGAGGTTTAAATCAACAGAAAACACAGAAATTGATCATCAT 2954
DB 215 TGGCTATTTAACCAATGAGAGGTTTAAATCAACAGAAAACACAGAAATTGATCATCAT 156
QY 2955 TTTGATACCTGCCATGTAACATCTACTCTCTGAAAATAAATGTTGGTCCATGCAGGGT 3012
DB 155 TTTGATACCTGCCATGTAACATCTACTCTCTGAAAATAAATGTTGGTCCATGCAGGGT 98

RESULT 54
AX874760/c
LOCUS AX874760 587 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 9665 from Patent EP1074617.
ACCESSION AX874760
VERSION AX874760.1 GI:40029519
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1
AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
TITLE Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
JOURNAL Primers for synthesizing full-length cDNA and their use
Patent: EP 1074617-A 9665 07-FEB-2001;
Research Association for Biotechnology (JP)
Location/Qualifiers
1..587
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

FEATURES
source
Location/Qualifiers
1..587
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 6.0%; Score 188; DB 2; Length 308;
Best Local Similarity 99.7%; Pred. No. 5.4e-98;
Matches 308; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2586 CTTAGACTGTGGCCATGCAAGCGAGAAAAGTTCCCTCTGAAACCAATGTTTACTGCT 2645
DB 1 CTTAGACTGTGGCCATGCAAGCGAGAAAAGTTCCCTCTGAAACCAATGTTTACTGCT 60
```

```
QY 2646 CTTACATGGTTCTCGGATGAGATGCCAATTTTGACATATTAATCTAGTTT 2705
Db |||||
QY 61 CTTACATGGTTCTCGGATGAGATGCCAATTTTGACATATTAATCTAGTTT 120
Db |||||
QY 2706 TTTAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCAT 2765
Db |||||
QY 121 TTTAGTGAGGCTGGAAG-CAATATGATTTACAGATCTATCTCTCAGGAGACACAGCAT 179
Db |||||
QY 2766 AAGAGTTCTGAACTCGGAGAACATTAATGAACTGCACTTTTGGCACTACCTTCAAGAAAA 2825
Db |||||
QY 180 AAGAGTTCTGAACTCGGAGAACATTAATGAACTGCACTTTTGGCACTACCTTCAAGAAAA 239
Db |||||
QY 2826 CTTGGATCAGTATTCCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 2885
Db |||||
QY 240 CTTGGATCAGTATTCCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 299
QY 2886 GGTATACAC 2894
Db |||||
QY 300 GGTATACAC 308

RESULT 55
BD109703
LOCUS EST and encoded human protein. 308 bp DNA linear PAT 18-SEP-2002
DEFINITION
ACCESSION BD109703
VERSION BD109703.1 GI:23204521
KEYWORDS JP 2002010789-A/1780.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo;
1 (bases 1 to 308)
AUTHORS Edwards, J.B.D.M., Jobert, S. and Giordano, J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 1780 15-JAN-2002;
GENSET CORP
COMMENT OS Homo sapiens (human)
PN JP 2002010789-A/1780
PD 15-JAN-2002
PF 07-AUG-2006 JP 2002080989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC Von Heijne matrix
CC score 4.5999990463257
CC seq AnrSILSLFLVRA/GK
FH Key Location/Qualifiers
FT CDS 17..178
FT sig_peptide 17..133.
FEATURES
source Location/Qualifiers
1..308
/organism="Homo sapiens"
/db_xref="taxon:9606"
ORIGIN
Query Match 6.0%; Score 188; DB 2; Length 308;
Best Local Similarity 99.7%; Pred. No. 5.4e-98;
Matches 308; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 2586 CTTAGGATCTGTGGCCATGCAAGCAGAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCT 2645
Db 1 CTTAGGATCTGTGGCCATGCAAGCAGAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCT 60
QY 2646 CTTACATGGTTCTCGGATGAGATGCCAATTTTGACATATTAATCTAGTTT 2705
Db |||||
QY 61 CTTACATGGTTCTCGGATGAGATGCCAATTTTGACATATTAATCTAGTTT 120
Db |||||
```

```
QY 2706 TTTAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCAT 2765
Db |||||
QY 121 TTTAGTGAGGCTGGAAG-CAATATGATTTACAGATCTATCTCTCAGGAGACACAGCAT 179
Db |||||
QY 2766 AAGAGTTCTGAACTCGGAGAACATTAATGAACTGCACTTTTGGCACTACCTTCAAGAAAA 2825
Db |||||
QY 180 AAGAGTTCTGAACTCGGAGAACATTAATGAACTGCACTTTTGGCACTACCTTCAAGAAAA 239
Db |||||
QY 2826 CTTGGATCAGTATTCCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 2885
Db |||||
QY 240 CTTGGATCAGTATTCCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 299
QY 2886 GGTATACAC 2894
Db |||||
QY 300 GGTATACAC 308

RESULT 56
AR414150
LOCUS Sequence 1787 from patent US 6639063. 308 bp DNA linear PAT 18-DEC-2003
DEFINITION
ACCESSION AR414150
VERSION AR414150.1 GI:40169260
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 308)
AUTHORS Edwards, J.-B.D.M., Jobert, S. and Giordano, J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 1787 28-OCT-2003;
Genset S.A.;
WOX;
FEATURES
source Location/Qualifiers
1..308
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 6.0%; Score 188; DB 2; Length 308;
Best Local Similarity 99.7%; Pred. No. 5.4e-98;
Matches 308; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 2586 CTTAGGATCTGTGGCCATGCAAGCAGAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCT 2645
Db 1 CTTAGGATCTGTGGCCATGCAAGCAGAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCT 60
QY 2646 CTTACATGGTTCTCGGATGAGATGCCAATTTTGACATATTAATCTAGTTT 2705
Db 61 CTTACATGGTTCTCGGATGAGATGCCAATTTTGACATATTAATCTAGTTT 120
QY 2706 TTTAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCAT 2765
Db 121 TTTAGTGAGGCTGGAAG-CAATATGATTTACAGATCTATCTCTCAGGAGACACAGCAT 179
QY 2766 AAGAGTTCTGAACTCGGAGAACATTAATGAACTGCACTTTTGGCACTACCTTCAAGAAAA 2825
Db 180 AAGAGTTCTGAACTCGGAGAACATTAATGAACTGCACTTTTGGCACTACCTTCAAGAAAA 239
QY 2826 CTTGGATCAGTATTCCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 2885
Db 240 CTTGGATCAGTATTCCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 299
QY 2886 GGTATACAC 2894
Db |||||
QY 300 GGTATACAC 308

RESULT 57
AX522584
LOCUS Sequence 254 from Patent WO02064731. 168 bp DNA linear PAT 24-OCT-2002
DEFINITION
```

ACCESSION AX522584
VERSION AX522584.1 GI:24411538
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1. Telerman, A., Anson, R., Tuijinder, M. and Susini, L.
Sequences involved in phenomena of tumour suppression, tumour
reversion, apoptosis and/or virus resistance and their use as
medicines
JOURNAL Patent: WO 02064731-A 254 22-AUG-2002;
Molecular Engines Laboratories (FR)
FEATURES
source
1..168
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 5.4%; Score 167; DB 2; Length 168;
Best Local Similarity 100.0%; Fred. No. 1.3e-85;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2944 GATCATCACATTTTGATACCTGCCATGTAAACATCTACTCTCTGAAAAATAAATGTGTGCCA 3003
DB 168 GATCATCACATTTTGATACCTGCCATGTAAACATCTACTCTCTGAAAAATAAATGTGTGCCA 109
QY 3004 TGCAGGGGCTACGGTTTGTGTAGTAATCTTAATCACTTAACCCCATGCTCAAAATCA 3063
DB 108 TGCAGGGGCTACGGTTTGTGTAGTAATCTTAATCACTTAACCCCATGCTCAAAATCA 49
QY 3064 AATGATACATATTTCTGAGACCCAGCAATACCAATAGAAATTACTA 3110
DB 48 AATGATACATATTTCTGAGACCCAGCAATACCAATAGAAATTACTA 2
RESULT 58
AC068507/c
LOCUS
DEFINITION Homo sapiens chromosome 15 clone RP11-349H15 map 15, WORKING DRAFT
SEQUENCE, 27 unordered pieces.
ACCESSION AC068507
VERSION AC068507.2 GI:8389522
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 176282)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 15, clone RP11-349H15
Unpublished
2 (bases 1 to 176282)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bida, F.,
Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collins, A., Cooke, K., DeArelano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
Meldrum, J., Menes, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
Murphy, T., Navloz, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neil, D., Olivari, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,

TITLE
JOURNAL
COMMENT

Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tessaye, S., Theodre, J., Tirrell, A., Travers, M., Trigglio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (03-MAY-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 9, 2000 this sequence version replaced gi:7684421.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997),
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: Li0275
Center clone name: 349_H_15
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 156826 bases at least Q40
Consensus quality: 166608 bases at least Q30
Consensus quality: 170674 bases at least Q20
Insert size: 173682; sum-of-contigs
Quality coverage: 3.6 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1477: contig of 1477 bp in length
* 1478: gap of 100 bp
* 1578: contig of 1208 bp in length
* 2786: gap of 100 bp
* 2886: contig of 1974 bp in length
* 4860: gap of 100 bp
* 4960: contig of 1457 bp in length
* 6417: gap of 100 bp
* 6517: contig of 2223 bp in length
* 8740: gap of 100 bp
* 8840: contig of 2029 bp in length
* 10869: gap of 100 bp
* 10969: contig of 2769 bp in length
* 13738: gap of 100 bp
* 13838: contig of 1492 bp in length
* 15329: gap of 100 bp
* 15429: contig of 2509 bp in length
* 17938: gap of 100 bp
* 18039: contig of 1943 bp in length
* 18039: gap of 100 bp
* 19982: contig of 3506 bp in length
* 20082: gap of 100 bp
* 23588: contig of 3760 bp in length
* 27448: gap of 100 bp
* 27548: contig of 2614 bp in length
* 30162: gap of 100 bp
* 30262: contig of 4476 bp in length
* 34738: gap of 100 bp
* 34838: contig of 5264 bp in length
* 40102: gap of 100 bp
* 40202: contig of 3838 bp in length
* 44039: gap of 100 bp
* 44140: contig of 5811 bp in length
* 49551: gap of 100 bp
* 50505: contig of 5505 bp in length
* 55555: contig of 5505 bp in length

Thu Jun 22 09:04:25 2006

us-10-825-632-2.olg.rge

```
/db_xref="taxon:9606"
ORIGIN
Query Match      2.6%; Score 82; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.9e-35;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3039 CCTTAACCCCATGCTCAAAATCAAATGATACATATTCCTGAGAGACCCAGCAATACCA 3098
      |||
      1 CCTTAACCCCATGCTCAAAATCAAATGATACATATTCCTGAGAGACCCAGCAATACCA 60
      |||

Db 3099 TAAGAATTACTAAAAA 3120
      |||
      61 TAAGAATTACTAAAAA 82
      |||

RESULT 62
BV448370 477 bp DNA linear STS 30-MAR-2005
LOCUS BV448370 Rhesus macaque genomic DNA Macaca mulatta STS genomic
DEFINITION clone WMA9839, sequence tagged site.
ACCESSION BV448370.1 GI:62000158
VERSION BV448370.1
KEYWORDS STS.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Cercopithecoidea; Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 477)
AUTHORS Spindel,E.R., Pauley,M., Jia,Y., Thompson,S., Lankhorst,M.,
          Gravett,C., Lupo,S.L., Tchourbanov,A., Ali,H., Ojeda,S.R. and
          Norgren,R.B.
          Targeted amplification of the 3' end of rhesus macaque orthologs of
          human genes
          Unpublished (2004)
JOURNAL
COMMENT
Contact: Spindel ER
Division of Neuroscience
Oregon National Primate Research Center
505 NW 185th Avenue, Beaverton, OR 97006, USA
Tel: 403-690-5388
Fax: 503-690-5384
Email: spindel@ohsu.edu
Primer A: agaacaattatgaactgcattct
Primer B: tttaagtgtatcaaatctgt
STS size: 477
PCR Profile:
  Hot Start: 95 degrees C for 2.00 min
  Denaturation: 95 degrees C for 0.50 min
  Annealing: 51 degrees C for 0.50 min
  Polymerization: 72 degrees C for 1.00 min
  PCR Cycles: 35
  Extension: 72 degrees C for 7.0 min
  Thermal Cycler: MJ Instruments PTC100
Protocol:
  Template: 200 ng
  Primer: each 1uM
  dNTP's: each 200 uM
  Tag Polymerase: 0.05 units/ul (Fast Start High
  Fidelity, Roche)
  Total Vol: 50 ul
Buffer:
  MgCl2: 1.8 mM
  Fast Start polymerase reaction buffer (Roche)

Bases 1-320 are 97% homologous (Blast) to bases 2805-3124 of
NM_197961.1. Primers were chosen to amplify genomic DNA in the 3'
region of DPP8. As human sequence was used to design the primers,
the primer sequences are not included in the rhesus sequence
provided below. To obtain additional information regarding primers
or clones contact: Dr. Robert Norgren; Dept of Genetics, Cell
```

Biolog & Anatomy: University of Nebraska Medical Center; 986395
 Nebraska Medical Center; Omaha, NE 68198. Email:
 rnoergren@unmc.edu
 A database containing sequences associated with this project can be
 found at: <http://thegenechip.unomaha.edu/index.html>.

FEATURES

source

```
1..477
  /organism="Macaca mulatta"
  /mol_type="genomic DNA"
  /strain="Indian origin"
  /db_xref="taxon:9544"
  /clone_MIM="9839"
  /clone_lib="Rhesus"
  /dev_stage="Adult"
  /notes="Organ: Liver; Vector: pGEM-T Easy; V-type: Plasmid;
  S1S was amplified from rhesus genomic DNA with the human
  forward and reverse primers listed above and subcloned
  into pGEM-T Easy"
1..477
  /genes="DPP8"
  /notes="dipeptidylpeptidase 8"
<1..5477
  /genes="DPP8"
```

gene

STS

ORIGIN

Query Match 2.6%; Score 82; DB 7; Length 477;
 Best Local Similarity 100.0%; Pred. No. 2e-35;
 Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3018 GTTTGTGGTAGTAACTTAATACCTTAACCCACATGCTCAAAATCAATGATCATATTC 3077
 |||||
 Db 214 GTTTGTGGTAGTAACTTAATACCTTAACCCACATGCTCAAAATCAATGATCATATTC 273
 |||||

Qy 3078 CTGAGAGACCCAGCAATACCAT 3099
 |||||
 Db 274 CTGAGAGACCCAGCAATACCAT 295
 |||||

RESULT 63

AC157219

LOCUS

DEFINITION

AC157219.2

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AC157219 222034 bp DNA linear HTG 01-JUL-2005
 Bos taurus clone CH240-74L16, *** SEQUENCING IN PROGRESS ***, 25
 unordered pieces.

AC157219 GI:68301606
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

Bos taurus (cattle)
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.

1 (bases 1 to 222034)
 Muzny, D. Marie, Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Bielawski, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Bunay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

Karpavich, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowals, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenuheva, L., Louised, H., Lozano, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindran, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapus, P., Martin, K., Martin, R., Martinez, E.,
 Mawhinney, S., McLeod, M.E., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, N., Munidasa, M., Murphy, M., Nair, L.,
 Nwakolaleh, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwankwelu, O., Okwuon, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Ploner, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
 Puzos, J., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Kelly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Riggs, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, M., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
 Smeed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, D., Tingey, A., Trejos, Z., Umani, K.,
 Velasquez, V., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wlaczek, R., Wood, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

2 (bases 1 to 222034)

Direct Submission

Worley, K.C.

Submitted (10-FEB-2005) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 222034)

Direct Submission

Cow Genome Sequencing Consortium.

Submitted (01-JUL-2005) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Jun 29, 2005 this sequence version replaced gi:58866073.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described

in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold') within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with 'N's to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature

table

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: FECE

Center clone name: CH240-74L16

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 216001 bases at least Q40

Consensus quality: 217476 bases at least Q30

Consensus quality: 218621 bases at least Q20

Estimated insert size: 219135; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently consists of 25 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 29449: contig of 29449 bp in length
29450 29615: gap of 166 bp
35158: contig of 5543 bp in length
35159 35208: gap of 50 bp
35209 46041: contig of 10833 bp in length
46042 46091: gap of 50 bp
46092 65721: contig of 19630 bp in length
65722 65771: gap of 50 bp
70033: contig of 4262 bp in length
70034 70083: gap of 50 bp
70084 75270: contig of 5187 bp in length
75271 75459: gap of 189 bp
75460 77510: contig of 2051 bp in length
77511 77739: gap of 229 bp
77740 87986: contig of 10247 bp in length
87987 88036: gap of 50 bp
88037 89277: contig of 1891 bp in length
89278 89977: gap of 50 bp
89978 92059: contig of 2082 bp in length
92059 92109: gap of 50 bp
92110 93888: contig of 1779 bp in length
93889 93938: gap of 50 bp
93939 98456: contig of 4518 bp in length
98457 98506: gap of 50 bp
98507 115454: contig of 16948 bp in length
115455 115504: gap of 50 bp
115505 116861: contig of 1357 bp in length
116862 116911: gap of 50 bp
116912 129154: contig of 12243 bp in length
129155 129204: gap of 50 bp
129205 137024: contig of 7820 bp in length
137025 137074: gap of 50 bp
137075 156249: contig of 19175 bp in length
156250 156299: gap of 50 bp
156300 159047: gap of 50 bp
159047 159097: contig of 6764 bp in length
159097 165910: gap of 50 bp
165911 178253: contig of 12343 bp in length
178254 178303: gap of 50 bp
178304 180139: contig of 1896 bp in length
180140 180249: gap of 50 bp
180250 217650: contig of 37401 bp in length
217651 217750: gap of unknown length
217751 218931: contig of 1181 bp in length
218932 219031: gap of unknown length
219032 220396: contig of 1365 bp in length
220397 220497: gap of unknown length
220497 222034: contig of 1538 bp in length.

FEATURES

source

1. 222034
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-74U16"
29450..29615
/estimated_length=166
35159..35208
/estimated_length=50
46042..46091
/estimated_length=50
65722..65771
/estimated_length=50
70034..70083
/estimated_length=50

gap
gap
gap
gap
gap

gap
gap
gap
gap
gap
gap
gap
gap
gap

75271..75459
/estimated_length=189
77511..77739
/estimated_length=229
87987..88036
/estimated_length=50
89928..89977
/estimated_length=50
92050..92109
/estimated_length=50
93889..93938
/estimated_length=50
98457..98506
/estimated_length=50
115455..115504
/estimated_length=50
116862..116911
/estimated_length=50
129155..129204
/estimated_length=50
137025..137074

Query Match 2.6%; Score 81; DB 12; Length 222034;
Best Local Similarity 100.0%; Pred. No. 8.8e-35;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1340 CCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTTCATCTACCTGAATTAT 1399
Db 118848 CCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTTCATCTACCTGAATTAT 118907
Qy 1400 TTATCCAGTAGAGATGATG 1420
Db 118908 TTATCCAGTAGAGATGATG 118928

RESULT 64

AC106509/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

224227 bp DNA linear HTG 13-MAY-2003
Rattus norvegicus clone CH230-20218, WORKING DRAFT SEQUENCE, 3
unordered pieces.
AC106509 4 GI:30580817
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus
Rattus norvegicus
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 224227)
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleaveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davison,S., Deramo,C., Ding,Y., Dinh,H., Diya,K.,
Delgado,O., Denson,S., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregiorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hollins,B., Howells,S., Hults,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J., and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 4799)
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
Thomas L. Casavant.
Web site: <http://genome.uiowa.edu>
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,
Fishler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,
Schneetz,T., Smith,C., Shier,E., Tack,D., Trout,K., Walters,J.,
Casavant,T., Soares,M.B.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: Plate: Row: Column: 0
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 27229207.
Location/Qualifiers
1. .4799
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:58057 IMAGE:6410075"
/tissue_type="Brain, enriched mouse brain 12.5dp"
/clone_lib="NIH BMAP_F00"
/lab_host="DH10B"
/note="Vector: pYX-ASC"
1. .4799
/gene="Dpp8"
/db_xref="GeneID:74388"
/db_xref="MGI:1921638"
288. .2966
/gene="Dpp8"
/codon_start=1
/product="dipeptidylpeptidase 8"
/protein_id="AA43124.1"
/db_xref="GI:27695450"
/db_xref="GeneID:74388"
/db_xref="MGI:1921638"
/translations="MKIPSGRCNMAAAMETEQLGVEIFETAECEGNGESQDRPKLEP
FYVERYSWLSQKLLADTRKYHGYMMKAPHPFMFKRTDPDPSHSDRVYLAWSGEN
RENTLYPSRTPKTINRAAVLMSKPLLDLFOATLDYGMYSREELLERKRIGTVGI
AANDYHPSGTFLPQAGSGIYHKDGGPHGFTQQLRNLVETSPNRPMDPKLCRAD
PDWATFHSNDIWSNLVTEERRITYVHNLNMEEDPRSGAVTFVQGEFDRKSG
YVMQFAERTPSGKKILRLIDVLEENDESEVEIHVTSPLMETRADSPRYKPTANPK
VTFKMSIVVDAAGGIDVIDKELQVPEILFEGVEYIARAGTPEGHAWSLIDRS
QTHLQVLIISPELIPVEDDAMDRQLIESVDSVPTLIIEETDIDWINHIDIFHVF
POTHEDEIFETASECKTGRHLKYITKLSLKESKYSKSGGLPAPSPDFCKIEEITI
TSGEVFLGRHGSNIWDEARKLVYFEGTKDPLHHLVVTSYANPGEVRLTDRGYS
HSCSLSDCFDFTSKYSNOKNPHCVSLYKLSPEDDPVHKTKEFWATILDAGPLPDY
TPPFIFFESTITDYNRGSCHGLKFGKPFKYMGOIIDDQVGLQYLASQYDFIDL
INTLASIGYVVVVVDNRGSKGLKFGKPFKYMGOIIDDQVGLQYLASQYDFIDL
DRVGIHWSYGGYISLMALMQRSIDIFRAIAGAPVTLWIFDTGYTRYMGHPDQNEQ
GYVLGSVAMQAEKFPSPBNPRLLLHLGFLDENVHFAHTSILLFLVRAGKPYDLQIYPQ

ORIGIN
Query Match 1.8%; Score 56; DB 6; Length 4799;
Best Local Similarity 100.0%; Pred. No. 4.6e-20;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1999 AAAACAAGGAATTTGGCCACCAATTTGGATTGATTCAGCAGTCTCTTCTGACTA 2054
|||||
Db 2103 AAAACAAGGAATTTGGCCACCAATTTGGATTGATTCAGCAGTCTCTTCTGACTA 2158
|||||
RESULT 66
LOCUS BC059222 4799 bp mRNA linear ROD 30-JUN-2004
DEFINITION Mus musculus dipeptidylpeptidase 8, mRNA (cdna clone MGC:66620
IMAGE:6410075), complete cds.
ACCESSION BC059222.1 GI:37590653
VERSION BC059222
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Strausberg,R.B., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,J., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heideh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schneetz,T.E., Brownstein,M.J., Udwin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J., and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 4799)
Strausberg,R.
Direct Submission
Submitted (01-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 125 Row: i Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 27229207.

FEATURES
source

Location/Qualifiers
1. .4799
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:66620 IMAGE:6410075"
/tissue_type="brain, enriched mouse brain 12.5dp"
/clone_lib="NIH BMAP_F00"
/lab_host="DH10B"
/note="vector: pYX-ASC"

gene

1. 4799
/gene="Dpp8"
/db_xref="GeneID:74388"
/db_xref="WGI:1921638"
288..2966
/gene="Dpp8"
/codon_start=1
/product="dipeptidylpeptidase 8"
/protein_id="AAH59222.1"
/db_xref="GI:137590654"
/db_xref="GeneID:74388"
/db_xref="WGI:1921638"

CDS

translation="MKIPGRCNMAAMETQLGVEIFETACEEGNGESQDRPKLEP
FYRYSWQLKKLLADTRKCHGYMAKAPHDPMFKRTDPDGPHSDRVVYLAWSGEN
KENTLYSEPTKINRAAGIYVLMISWPKLLDLQAATLDYGMVREELRLRKRIGTGTG
AAYDHPGSGTFLPAGSGIYHIKDGPHGFTQQLPRLNVLVETSCPNRMWPKLCAD
PMWIAFHNDNLWNLVREERTIYVHNLANNEEDPRSGAVATPVLQOEFRYSG
YMWCPQAEPTSGGKILRLILENDESEVEIHTVSPMLETRADSPVFKTSTANPK
VTPKMSIYVDAAGGIIDVDELVQPEILFEGVEVTLIAREAGMTPEGKHAWSLLDERS
QTHQLVLSPLEFLPVEDDAMDRLIESVPDSVTLIAREAGMTPEGKHAWSLLDERS
TQHEDEIFPASECKTGFRHLVKITSLKESYKRSSGGLPAPDPKCPKIKETI
TQHEVLRGCHNINWDEARKLVYFEGTKDSPLEHLVYTSVANPEVPLTDRGYS
HSCCLSRHCDFFISYKNQPHCVSLYKLSPPEDDPVHKTFKFNATLDSAGPLDY
TPPEIFSTFTGLYQMLYKPHLOPKKYPTVLPYIGYQVQLVNNRFGVKYKFR
LNTLASLGVVVDINRSGCHRGKLFEGAPKYKQLEIIDDVEGLQYLASQDFIDL
DRVGLHSGYVVDINRSGCHRGKLFEGAPKYKQLEIIDDVEGLQYLASQDFIDL
GYLGSVMAQAEKFPSEPNRLLLHGLDENVHAHTSILLSFLVRAGAPYDLQIYPO
ERHSRVAPSGEYHLLHVLQENLSRAALKYI"

ORIGIN

Query Match 1.8%; Score 56; DB 6; Length 4799;
Best Local Similarity 100.0%; Pred. No. 4.6e-20;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1999 AAAACAAGGAATTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTA 2054
|||||
Db 2103 AAAACAAGGAATTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTA 2158
|||||

RESULT 67

AC1414143/c
LOCUS AC1414143 305843 bp DNA linear HTG 09-APR-2003
DEFINITION Macaca mulatta clone CH250-272H19, *** SEQUENCING IN PROGRESS ***.
ACCESSION AC1414143
VERSION AC1414143.1 GI:29649813
KEYWORDS HTG; HTGS PHASE2; HTGS_FGI.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecoidea; Macaca.
1. (bases 1 to 305843)
Cauros, M. and Milosavljevic, A.
Pooled genomic indexing (PGI): mathematical analysis and experiment
design
(in) Guigo, R. and Gusfield, D. (Eds.);
ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI
2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-28;
Springer (2002)
2. (bases 1 to 305843)
Milosavljevic, A., Sodergren, E., Cauros, M., Li, B., Jackson, A.R.,
Adams, C., Adio-Oduola, B., Ali-Osman, P.R., Allen, C., Alsbrooks, S.L.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission
Submitted (09-APR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center of Medicine
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
Project information
Center project name: LCRF
Center clone name: CH250-272H19
----- Summary Statistics
Chemistry: Dye-terminator Big Dye; Inf% of reads
Chemistry: Dye-terminator Big Dye; Inf% of reads
Consensus quality: 12640 bases at least Q40
Consensus quality: 15085 bases at least Q30
Consensus quality: 17385 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* NOTE: The contigs are based on the application (NCBI build 31)
* of the PGI method using the Human genome (NCBI build 31)
* as the comparative genome.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have

Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbara, J.,
Benton, J., Bivaga, K., Blankenburg, K., Bonnin, D., Bouck, J.,
Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, R.,
Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D.,
Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
Davy, Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,
Dugan-Rocha, S., Durbin, K.J., Egan, A., Farnhart, C., Edwards, C.C.,
Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D.,
Flagg, A., Garner, T., Foster, F., Frantz, P., Gabisi, A., Gao, J.,
Garcia, N., Hale, S., Hamilton, K., Han, J., Harris, C., Guevara, W.,
Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Guevara, W.,
Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, K.,
Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Honsi, F.,
Howard, S., Hu, Y., Hui, S., Hume, J., Ioshikhes, I., Jackson, L.E.,
Jacobson, B., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Karlsson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C.,
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
Luisseg, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R.,
Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P.,
Meador, M., Mei, G., Merscher, S., Metzker, M., Miller, A., Miner, G.,
Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M.,
Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N.,
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokwenkwo, S.,
Oguy, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
Perez, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.,
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M.,
Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shim, C.,
Shoostari, N., Sisson, I., Sodergren, E., Sonaie, T., Sparks, A.,
Stanley, H., Stone, H., Sutton, A., Svatek, A., Taber, P., Tamerisa, A.,
Tamerisa, K., Tansey, J., Taylor, C., Taylor, T., Telford, B.,
Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D.,
Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R.,
Washington, C., Watlington, S., Williams, G., Williams, A., Zhou, J.,
Wlezyka, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y., Zhou, J.,
Zorilla, S., Zscherpel, R., Weinstein, G. and Gibbs, R.

* provided by the submittor.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 305843: contig of 305843 bp in length.

```

FEATURES
    source
        1..305843
            /organism="Macaca mulatta"
            /mol_type="genomic DNA"
            /db_xref="taxon:9544"
            /clone="CH250-272H19"
    misc_feature
        1..305843
            /note="assembly name:CH250-272H19.1B
            CONFIDENCE: 0.83"

ORIGIN
    Query Match          1.7%; Score 54; DB 12; Length 305843;
    Best Local Similarity 100.0%; Pred. NO. 7.6e-19;
    Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1167 AGGTACAGCAATCTTAAGTCACCTTTAAGATGTCAGAAATATGATTGATGC 1220
      |||||
DB 131782 AGGTACAGCAATCTTAAGTCACCTTTAAGATGTCAGAAATATGATTGATGC 131729

RESULT 68
AC144143 305843 bp DNA linear HTG 09-APR-2003
LOCUS Macaca mulatta clone CH250-272H19, *** SEQUENCING IN PROGRESS ***.
AC144143
DEFINITION AC144143.1 GI:29649813
VERSION HTG; HTGS PHASE2; HTGS PGI.
KEYWORDS Macaca mulatta (rhesus monkey)
SOURCE Macaca mulatta
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecinae; Macaca.

REFERENCE
1 (bases 1 to 305843)
Cauros,M. and Milosavljevic,A.
Pooled genomic indexing (PGI): mathematical analysis and experiment
design
(in) Guigo,R. and Gusfield,D. (Eds.);
ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI
2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-28;
Springer (2002)

2 (bases 1 to 305843)
Milosavljevic,A., Sodergren,E., Cauros,M., Li,B., Jackson,A.R.,
Adams,C., Adio-Oduola,B., Ali-ouman,F.R., Allen,C., Alsbrooks,S.L.,
Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbacia,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,
Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Deigado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Eugan-Rocha,S., Durbin,K.J., Egan,A., Earnhart,C., Edwards,C.C.,
Dulanja-C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,
Elhaj,C., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,
Flagg,N., Ford,J., Foster,P., Frantz,P., Gill,R., Gorrell,J.H., Guevara,M.,
Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Harris,K.,
Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Hernandez,O.,
Hart,M., Haviak,P., Hawes,A., Hernandez,J., Hollins,B., Homsai,F.,
Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Ioshikhes,I., Jackson,L.E.,
Howard,S., Huber,J., Hulyk,S., Hume,J., Jolivet,S., Joudah,S.,
Jacobson,B., Jia,Y., Johnson,R., Johnson,U., King,L., Korvah,J., Kovar,C.,
Karleson,E., Kelly,S., Khan,U., King,L., Lee,B., Lee,E., Lewis,L.C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Liu,W.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Lucier,R., Luna,R.,
Louiseged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Martin,R.,
Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martini,R.,
Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P.,
Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G.,
  
```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: LCRF
Center clone name: CH250-272H19
----- Summary Statistics
Chemistry: Dye-terminator Big Dye: inf% of reads
Chemistry: Dye-terminator Big Dye: inf% of reads
Consensus quality: 12640 bases at least Q40
Consensus quality: 15085 bases at least Q30
Consensus quality: 17385 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: The contigs are based on the application
* of the PGI method using the Human genome (NCBI build 31)
* as the comparative genome.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 305843: contig of 305843 bp in length.
* Location/Qualifiers
    1..305843
        /organism="Macaca mulatta"
        /mol_type="genomic DNA"
        /db_xref="taxon:9544"
        /clone="CH250-272H19"
    misc_feature
        1..305843
            /note="assembly name:CH250-272H19.1B
            CONFIDENCE: 0.83"

ORIGIN
    Query Match          1.5%; Score 48; DB 12; Length 305843;
    Best Local Similarity 100.0%; Pred. No. 2.6e-15;
    Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 978 TGTTCTCCAGAGAAGATTTGATAGATATTCCTGGCTATTGGTGGTCTCC 1025
      |||||
DB 249937 TGTTCTCCAGAGAAGATTTGATAGATATTCCTGGCTATTGGTGGTCTCC 249984
  
```

RESULT 69

AC161443
 LOCUS Mus musculus chromosome 1, clone RP23-404K2, complete sequence.
 DEFINITION AC161443
 VERSION AC161443.14 GI:82177432
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 186144)

Birken, B., Nusbaum, C. and Lander, E.

Mus musculus chromosome 1, clone RP23-404K2

Unpublished

2 (bases 1 to 186144)

Birken, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
 Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J.,
 Cheopel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B.,
 DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
 Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,
 Galagan, J., Gardyna, S., Graham, L., Horton, L., Hulme, W., Hafez, N.,
 Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
 Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
 MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
 McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Nicol, R., Norbu, C.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemбек, L.,
 Zimmer, A. and Zody, M.

Direct Submission

Submitted (14-MAY-2005) Broad Institute of MIT and Harvard, 320

Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 186144)

Birken, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
 Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J.,
 Cheopel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B.,
 DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
 Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,
 Galagan, J., Gardyna, S., Graham, L., Horton, L., Hulme, W., Hafez, N.,
 Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
 Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
 MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
 McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Nicol, R., Norbu, C.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemбек, L.,
 Zimmer, A. and Zody, M.

Direct Submission

Submitted (21-OCT-2005) Broad Institute of MIT and Harvard, 320

Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 186144)

Birken, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
 Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J.,
 Cheopel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B.,
 DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
 Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,

Galagan, J., Gardyna, S., Graham, L., Horton, L., Hulme, W., Hafez, N.,
 Hagoopian, D., Hago, B., Hall, J., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
 Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
 MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
 McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Nicol, R., Norbu, C.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemбек, L.,
 Zimmer, A. and Zody, M.

Direct Submission

Submitted (12-NOV-2005) Broad Institute of MIT and Harvard, 320

Charles Street, Cambridge, MA 02141, USA

On Nov 12, 2005 this sequence version replaced gi:77798131.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Broad Institute of MIT and Harvard

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L32392

Center clone name: 404_K_2

FEATURES

source

Location/Qualifiers
 1. 186144
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="1"
 /map="1"

/clone="RP23-404K2"
 /clone_lib="RPCI-23 Female Mouse BAC"

1. 272

repeat_region

/rpt_family="Llmd_F2"

repeat_region

/rpt_family="Llmd_F2"

repeat_region

complement(1428..1880)

repeat_region

/rpt_family="RLR17"

repeat_region

complement(2701..2775)

repeat_region

/rpt_family="Llmd_F2"

repeat_region

3013..3045

repeat_region

/rpt_family="HALib"

repeat_region

3046..3070

repeat_region

/rpt_family="(TTTTG)n"

repeat_region

3071..3133

repeat_region

/rpt_family="HALib"

repeat_region

complement(4980..5501)

repeat_region

/rpt_family="Lx2"

repeat_region

5691..5713

repeat_region

/rpt_family="(TTTTG)n"

repeat_region

6467..6493

repeat_region

/rpt_family="(A)n"

repeat_region

6541..6866

repeat_region

/rpt_family="ORRIA3"

repeat_region

7774..7807

repeat_region

/rpt_family="AT_Rich"

repeat_region

8071..8211

repeat_region

/rpt_family="BLF"

repeat_region

8656..8732

repeat_region

/rpt_family="(GA)n"

repeat_region

complement(8970..9532)

repeat_region

/rpt_family="RLR31_Mm"

repeat_region

9548..9602

repeat_region

/rpt_family="(TTTTTG)n"

repeat_region

complement(9603..9750)

[illegible]

1 (bases 1 to 197073)

Murny, D., Adams, C., Agbai II, O., Allen, C., Alsbrooks, S., Archer, P., Arredondo, H., Banderanaide, D., Bangura, L., Beltran, B., Beltran, P., Beraducci, A., Biswalo, K., Blyth, P., Bonham, H., Buhay, C., Burch, R., Cadoree, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I., Chacko, J., Chahrour, M., Chavez, D., Chen, A., Chen, G., Chen, R., Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A., Curry, S., Dai, W., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinu, H., Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, C., Dzieda, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M., Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P., Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T., Gaskin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D., Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haeblerlen, K., Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Haveliak, P., Haves, A., Hawkins, E., Hayes, S., Hernandez, J., Hines, S., Hitchens, M., Hodgson, A., Hogues, M., Hollins, B., Howell, L. T., Hulyk, S., Hume, J., Imo, K., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalatrus, K., Kelly, S., Keys, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Kara, F., Leal, S., Lee, K., Lee, S., Legall, F. I., Lemon, S., Lewis, L., Li, B., Li, Y., Li, Z., Linnell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D., London, P., Lopez, J., Lorenshewa, L., Lozado, R., Luk, T., Madu, R., Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E., McClelland, H., McPherson, J., Mercadado, C., Morris, S., Munidasa, M., Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Norwig-Eastaugh, E., Murray, D., Nazarith, L., Ngo, D., Nguyen, N., Norwig-Eastaugh, E., Nott, A., Nwaekemele, O., Oregon, M., Ochi-Okorie, C., Odeh, E., Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B., Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T., Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quiroz, J., Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S.,

[illegible]

Rivas, C., Rodriguez, P., Rojas, A., Ruiz, S.J., Sana, M., Sanders, W., Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, H., Shen, Y., Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thörn, R., Thornton, R., Trejos, Z., Umami, K., Vargo, C., Verduzco, D., Villaseña, D., Virk, D., Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R., Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L., Zhang, Z., Zhou, J., Weinstock, G. and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

CONSRM

TITLE

JOURNAL

COMMENT

Unpublished
2 (bases 1 to 197073)
Worley K.C.

Direct Submission

Submitted (11-NOV-2005)

Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 197073)

Bovine Genome Sequencing Consortium

Direct Submission

Submitted (25-JAN-2006) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Jan 25, 2006 this sequence version replaced gi:82174753. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: PRK1

Center clone name: CH240-528L18

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 190844 bases at least Q40

Consensus quality: 192231 bases at least Q30

Consensus quality: 193415 bases at least Q20

Estimated insert size: 192667; sum-of-contigs estimation

Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 21967: contig of 21967 bp in length

21968 22017: gap of 50 bp

22018 23050: contig of 1033 bp in length

23051 23159: gap of 109 bp

23160 24660: contig of 1501 bp in length

24661 24760: gap of unknown length

24761 26134: contig of 1374 bp in length

26135 26940: gap of 806 bp

26941 38125: contig of 11185 bp in length

38126 38175: gap of 50 bp

* 38176 48967: contig of 10692 bp in length
* 48968 48917: gap of 50 bp
* 48918 58217: contig of 9300 bp in length
* 58218 58267: gap of 50 bp
* 58268 160395: contig of 102128 bp in length
* 160396 161001: gap of 606 bp
* 161002 168377: contig of 7376 bp in length
* 168378 168427: gap of 50 bp
* 168428 179285: contig of 10857 bp in length
* 179286 179428: gap of 143 bp
* 179429 191688: contig of 12260 bp in length
* 191689 191787: gap of unknown length
* 191788 192959: contig of 1171 bp in length
* 192960 193058: gap of unknown length
* 193059 194632: contig of 1573 bp in length
* 194633 194731: gap of unknown length
* 194732 195949: contig of 1217 bp in length
* 195949 196048: gap of unknown length
* 196049 197073: contig of 1025 bp in length.

FEATURES

source

1..197073
/organism="Bos taurus"
/mol_type="Genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-528L18"
21968..22017
/estimated_length=50
23051..23159
/estimated_length=109
24661..24760
/estimated_length=unknown
26135..26940
/estimated_length=806
38126..38175
/estimated_length=50
48968..48917
/estimated_length=50
58218..58267
/estimated_length=50
160396..161001
/estimated_length=606
168378..168427
/estimated_length=50
179285..179427
/estimated_length=143
191688..191787
/estimated_length=unknown
192959..193058
/estimated_length=unknown
194632..194731
/estimated_length=unknown
195949..196048
/estimated_length=unknown

ORIGIN

Query Match 1.4%; Score 44; DB 12; Length 197073;
Best Local Similarity 100.0%; Pred. No. 6e-13;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 271 TGTGAGGAGATATTGATTCACAGATCGGCTAAATTGGAGCC 314
Db 22514 TGTGAGGAGATATTGATTCACAGATCGGCTAAATTGGAGCC 22557

RESULT 72

AC112162/c

LOCUS

DEFINITION

AC112162

ACCESSION

VERSION

KEYWORDS

SOURCE

AC112162
Mus musculus chromosome 9 BAC clone MGS1-296M6 ES cell line,
complete sequence.
AC112162
AC112162.8 GI:21327375
HTG.
Mus musculus (house mouse)

ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Mus.
REFERENCE AUTHORS	1 (bases 1 to 116580) Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zencneck,W., Xi.C., Juels,P. and Kucherlapati,R.
TITLE	High Throughput Mouse Sequencing
JOURNAL	Unpublished
REFERENCE AUTHORS	2 (bases 1 to 116580) Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zencneck,W., Xi.C., Juels,P. and Kucherlapati,R.
TITLE	Direct Submission
JOURNAL	Submitted (20-FEB-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA
REFERENCE AUTHORS	3 (bases 1 to 116580) Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zencneck,W., Xi.C., Juels,P. and Kucherlapati,R.
TITLE	Direct Submission
JOURNAL	Submitted (04-MAY-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA
REFERENCE AUTHORS	4 (bases 1 to 116580) Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zencneck,W., Xi.C., Juels,P. and Kucherlapati,R.
TITLE	Direct Submission
JOURNAL	Submitted (04-JUN-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA
REFERENCE AUTHORS	5 (bases 1 to 116580) Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zencneck,W., Xi.C., Juels,P. and Kucherlapati,R.
TITLE	Direct Submission
JOURNAL	Submitted (06-JUN-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA
REFERENCE AUTHORS	6 (bases 1 to 116580) Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zencneck,W., Xi.C., Juels,P. and Kucherlapati,R.
TITLE	Direct Submission
JOURNAL	Submitted (21-SEP-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA
COMMENT	On Jun 6, 2002 this sequence version replaced gi:21321767. -----Genome Center: Center: Harvard Partners Genome Center Center Code: HPGC Web site: http://www.hpcgg.org/Sequence/mouse.html Contact: hpgcmendel.mgh.harvard.edu
	CLONE LENGTH: This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the Features listing.
	ANNOTATION OF FEATURES: STSS are identified using ePCR (Genome Res. 7:541-550). Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences. Genes and Regions of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to ESTs and cDNA sequences in Unigene. Genes demonstrate at least two exons flanked by consensus splice sites that maintain sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

```

-----Summary Statistics
Center project name: ALI
Sequencing vector: pSMART; AF399742
Chemistry: Dye-terminator Big Dye; 100#
Assembly program: Phrap version 0.990319
Contig length: 116580
Fraction of Phrap value < 40: 0
Error Rate in Consed: 0.00 per 10,000 bases
Number of N's in consensus: 0

----- Distribution of Quality < 40 Bases:
                                     5      10      15      20      25      30      35      40
-----
# bases
1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
  0

-----
Location/Qualifiers
1. .116580
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="ES cell line"
/db_xref="taxon:10090"
/chromosome="9"
/clone="MGS1-296M6"
235. .369
/standard_name="D9Mit316"
repeat_region
258. .303
/rpt_family="(TG)n"
612. .1204
/standard_name="Nope"
complement(1584. .1643)
/rpt_family="B4A"
1710. .1748
/rpt_family="(TG)n"
complement(3793. .3909)
/rpt_family="PB1D10"
4042. .4191
/rpt_family="MIR"
4247. .4435
/rpt_family="B2_Mm2"
complement(4504. .4623)
/rpt_family="RSINE1"
4720. .4831
/rpt_family="L2"
complement(4840. .4879)
/rpt_family="MIR"
4938. .5118
/rpt_family="B2_Mm2"

```



```

repeat_region      6216..6250
                    /rpt_family="AT_rich"
repeat_region      complement(6251..6381)
                    /rpt_family="RSINE1"
repeat_region      complement(6421..6566)
                    /rpt_family="B1_MM"
repeat_region      6593..6612
                    /rpt_family="B1_MM"
repeat_region      /rpt_family="(T)n"
repeat_region      complement(6668..6790)
                    /rpt_family="B2_MM"
repeat_region      complement(6795..6820)
                    /rpt_family="B1_MM"
repeat_region      /rpt_family="B1_MM"
repeat_region      complement(6883..6954)
                    /rpt_family="B1_MM"
repeat_region      complement(6982..7058)
                    /rpt_family="B1_MM"
repeat_region      8050..8246
                    /rpt_family="B2_MM"
repeat_region      complement(8445..8719)
                    /rpt_family="B4"
repeat_region      complement(9019..9119)
                    /rpt_family="MIR"
repeat_region      9285..9371
                    /rpt_family="MIR"
repeat_region      complement(10773..10846)
                    /rpt_family="PB1D10"
repeat_region      10859..10956
                    /rpt_family="B4"
repeat_region      complement(11043..11141)
                    /rpt_family="PB1D7"
repeat_region      11848..11884
                    /rpt_family="(TC)n"
repeat_region      11884..11922
                    /rpt_family="(TC)n"
repeat_region      11930..11966
                    /rpt_family="(TC)n"
repeat_region      complement(11967..11994)
                    /rpt_family="B4"
repeat_region      complement(11995..12127)
                    /rpt_family="B1_MM"
repeat_region      complement(12128..12143)
                    /rpt_family="B4"
repeat_region      complement(12170..12252)
                    /rpt_family="RSINE1"
repeat_region      12853..12993
                    /rpt_family="B1_MM"
repeat_region      complement(13288..13367)

Query Match      1.3%; Score 42; DB 6; Length 116580;
Best Local Similarity 100.0%; Pred. NO. 9e-12;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      207 ATGCACATGCGACGACGATGGAAACAGACGCTGGGTGT 248
        |||||
Db      97296 ATGCACATGCGACGACGATGGAAACAGACGCTGGGTGT 97255

RESULT 73
AC161366/c
LOCUS
DEFINITION      Mus musculus BAC clone RP24-230D17 from chromosome 9, complete
sequence.
AC161366
VERSION      AC161366.6 GI:71533404
KEYWORDS      HTG.
SOURCE      Mus musculus (house mouse)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurgnath; Muridae; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 155733)
AUTHORS      Higginbotham,J., Kozlowicz,A. and Haglund,K.
TITLE      The sequence of Mus musculus BAC clone RP24-230D17

```

```

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

```

```

Unpublished (2001)
2 (bases 1 to 155733)
Wilson R.K.
Direct Submission
Submitted (12-MAY-2005) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 155733)
Wilson R.K.
Direct Submission
Submitted (17-JUN-2005) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 155733)
Wilson R.K.
Direct Submission
Submitted (30-JUL-2005) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 155733)
Wilson R.K.
Direct Submission
Submitted (17-AUG-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jul 30, 2005 this sequence version replaced gi:67514951.
----- Genome Center
Center: Washington University Genome Sequencing Center
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
----- Summary Statistics
Center project name: M_BB0230D17

```

COMMENT

NOTICE:

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e. phred quality >=30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence. The assembly was confirmed by restriction digest. This finishing standard has slightly changed from the previous Human standard. Specifically, standards for regions of low sequence complexity (such as dinucleotide repeats and small unit tandem repeats) have been relaxed. These regions are very prevalent in the mouse genome, and the return on extended finishing efforts is minimal. If a sequence meets the criteria of the above statement, it needs no comments or tags. If the criteria are not met, such as ambiguous bases, then the region is duly annotated.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The BAC library has been constructed by Pieter de Jong and coworkers (<http://www.choxi.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.choxi.org>

This sequence is the entire insert of the clone.

```

FEATURES
source

```

```

Location/Qualifiers
1..155733
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="9"
/clone="RP24-230D17"

```

unsaure /clone_lib="RPCI-24"
10180..10254
/notes="Sequence derived from one plasmid subclone."
51980..52027
/notes="Sequence derived from one plasmid subclone."
153091..153320
/note="Unresolved simple sequence repeat."

ORIGIN

Query Match 1.3%; Score 42; DB 6; Length 155733;
Best Local Similarity 100.0%; Pred.No.9.le-12; Indels 0; Gaps 0;
Matches 42; Conservative 0; Mismatches 0;

QY 1999 AAAAAAGGAATTTGGCCACCATTTTGGATTTCAGCAGGT 2040
|||||
DB 96576 AAAAAAGGAATTTGGCCACCATTTTGGATTTCAGCAGGT 96535
|||||

RESULT 74

AC122463 162528 bp DNA linear ROD 28-JAN-2005
Mus musculus BAC clone RP24-296C19 from 9, complete sequence.

AC122463

AC122463.4 GI:50839061

HTG

Mus musculus (house mouse)

SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

Tomlinson,C., Cotton,M. and Haglund,K.
1 (bases 1 to 162528)

AUTHORS

The sequence of Mus musculus BAC clone RP24-296C19

TITLE

Unpublished (2001)

REFERENCE

2 (bases 1 to 162528)

AUTHORS

McPherson,J.D. and Waterston,R.H.

TITLE

Direct Submission

JOURNAL

Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

REFERENCE

3 (bases 1 to 162528)

AUTHORS

Wilson,R.K.

TITLE

Direct Submission

JOURNAL

Submitted (10-JUL-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

REFERENCE

4 (bases 1 to 162528)

AUTHORS

Wilson,R.K.

TITLE

Direct Submission

JOURNAL

Submitted (30-JUL-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

REFERENCE

5 (bases 1 to 162528)

AUTHORS

Wilson,R.K.

TITLE

Direct Submission

JOURNAL

Submitted (28-JAN-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

COMMENT

On Jul 30, 2004 this sequence version replaced gi:50201506.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
----- Summary Statistics

Center project name: M_BB0296C19

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by

restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:

The RPCI-24 BAC library has been constructed by Pieter de Jong and
coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is
overlapped by AC140243.

FEATURES

source

1..162528
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="9"
/map="9"
/clone="RP24-296C19"
/clone_lib="RPCI-24"

repeat_region

932..1076

/rpt_family="Alu"

repeat_region

1787..1971

/rpt_family="MIR"

repeat_region

2617..2756

/rpt_family="MIR"

repeat_region

5649..5775

/rpt_family="Alu"

repeat_region

5776..5857

/rpt_family="B4"

repeat_region

7268..7672

/rpt_family="RMER17C"

repeat_region

8436..8519

/rpt_family="MIR"

repeat_region

8689..8757

/rpt_family="L2"

repeat_region

8817..8949

/rpt_family="Alu"

repeat_region

9301..9660

/rpt_family="MaLR"

repeat_region

10124..10303

/rpt_family="B2"

repeat_region

10355..10496

/rpt_family="B4"

repeat_region

10571..10718

/rpt_family="Alu"

repeat_region

10779..10908

/rpt_family="MaLR"

repeat_region

10906..11027

/rpt_family="B2"

repeat_region

11070..11222

/rpt_family="Alu"

repeat_region

11789..12045

/rpt_family="B4"

repeat_region

12241..12358

/rpt_family="B2"

repeat_region

12377..12536

/rpt_family="B4"

repeat_region

12593..12650

/rpt_family="L1"

repeat_region

12655..12800

/rpt_family="Alu"

repeat_region

12806..12922

/rpt_family="L1"

repeat_region

13137..13281

/rpt_family="Alu"

repeat_region

13289..13360

/rpt_family="Alu"

```
repeat_region /rpt_family="ID"
13532..13720 /rpt_family="B2"
repeat_region 14411..14550 /rpt_family="Alu"
repeat_region 17393..17539 /rpt_family="Alu"
repeat_region 17784..17977 /rpt_family="B2"
repeat_region 18370..18480 /rpt_family="B4"
repeat_region 18485..18614 /rpt_family="Alu"
repeat_region 19195..19364 /rpt_family="B2"
trna 19201..19273 /product="tRNA-Ser"
/notes="Likely pseudogene (HMM Sc=39.82 / Sec struct
Sc=11.21)"
repeat_region 19364..20105 /rpt_family="Alu"
repeat_region 20474..20688 /rpt_family="B2"
repeat_region 20770..20868 /rpt_family="Alu"
repeat_region 20960..21086 /rpt_family="MIR"
repeat_region 21223..21415 /rpt_family="B2"
repeat_region 21600..21746 /rpt_family="Alu"
repeat_region 21991..22060 /rpt_family="ID"
repeat_region 22845..22953 /rpt_family="Alu"
repeat_region 22965..23123 /rpt_family="B2"
repeat_region 23145..23338 /rpt_family="B2"
repeat_region 23760..23895 /rpt_family="B2"
repeat_region 23899..24191 /rpt_family="MaLR"
repeat_region 24476..24585 /rpt_family="B2"
repeat_region 24672..24818 /rpt_family="Alu"
repeat_region 25228..25408 /rpt_family="B2"
repeat_region 25503..25688 /rpt_family="B2"
repeat_region 25796..25880 /rpt_family="Alu"
repeat_region 25986..26135 /rpt_family="B4"
repeat_region 26233..26740 /rpt_family="Alu"
repeat_region 26750..26900 /rpt_family="Alu"
repeat_region 29766..29872 /rpt_family="Alu"
repeat_region 30031..30314 /rpt_family="B4"
repeat_region 31450..31578 /rpt_family="ID"
repeat_region 32188..32267 /rpt_family="B4"
repeat_region 32299..32426 /rpt_family="B4"
repeat_region 32479..32676 /rpt_family="B4"
repeat_region 32909..33004 /rpt_family="L1"
```

```
repeat_region 33598..33685
Query Match 1.3%; Score 42; DB 6; Length 162528;
Best Local Similarity 100.0%; Pred. No. 9.1e-12;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1999 AAAACAAGGAATTTTGGCCACCATTTTGGATTTCAGCAGGT 2040
DB 118998 AAAACAAGGAATTTTGGCCACCATTTTGGATTTCAGCAGGT 119039

RESULT 75
AR372392/c AR372392 612 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 31 from patent US 6395889.
ACCESSION AR372392
VERSION AR372392.1 GI:34609704
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 612)
TITLES Robison,K.E.
Nucleic acid molecules encoding human protease homologs
JOURNAL Patent: US 6395889-A 31 28-MAY-2002;
Millennium Pharmaceuticals, Inc.; Cambridge, MA
FEATURES Location/Qualifiers
source 1..612
/mol_type="genomic DNA"

Query Match 1.1%; Score 34; DB 2; Length 612;
Best Local Similarity 100.0%; Pred. No. 4.2e-07;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2495 CTGGGGCCCCAGTCACCTCTGTGATCTTCTATGA 2528
DB 365 CTGGGGCCCCAGTCACCTCTGTGATCTTCTATGA 332

RESULT 76
AC122776 72149 bp DNA linear HTG 25-MAY-2002
DEFINITION Mus musculus clone RP24-201E8, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC122776
VERSION AC122776.1 GI:21206389
KEYWORDS HTG; HTGS PHASE0.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 72149)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLES Mus musculus, clone RP24-201E8
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 72149)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Baota,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Charazak,B., Choepel,I., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., FitzHugh,W., Gage,D.,
Galegari,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Fierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Meneus,L., Minova,T., Menga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicoli,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
```

Thu Jun 22 09:04:25 2006

us-10-825-632-2.olig.rge

Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (25-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A. F. A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L26573
Center clone name: 201_E_8

* NOTE: This record contains 88 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 733: contig of 733 bp in length
* 833: gap of 100 bp
* 834 1531: contig of 718 bp in length
* 1552 1651: gap of 100 bp
* 1652 2370: contig of 719 bp in length
* 2371 2470: gap of 100 bp
* 2471 3207: contig of 737 bp in length
* 3208 3307: gap of 100 bp
* 3308 4023: contig of 716 bp in length
* 4024 4123: gap of 100 bp
* 4124 4833: contig of 710 bp in length
* 4834 4933: gap of 100 bp
* 4934 5636: contig of 703 bp in length
* 5637 5735: gap of 100 bp
* 5737 6411: contig of 675 bp in length
* 6412 6511: gap of 100 bp
* 6512 7248: contig of 737 bp in length
* 7249 7348: gap of 100 bp
* 7349 8088: contig of 740 bp in length
* 8089 8188: gap of 100 bp
* 8189 8919: contig of 731 bp in length
* 8920 9019: gap of 100 bp
* 9020 9758: contig of 739 bp in length
* 9759 9858: gap of 100 bp
* 9859 10588: contig of 730 bp in length
* 10589 10688: gap of 100 bp
* 10689 11420: contig of 732 bp in length
* 11421 11520: gap of 100 bp
* 11521 12236: contig of 716 bp in length
* 12237 12336: gap of 100 bp
* 12337 13033: contig of 697 bp in length
* 13034 13133: gap of 100 bp
* 13134 13866: contig of 733 bp in length
* 13867 13966: gap of 100 bp
* 13967 14670: contig of 704 bp in length
* 14671 14770: gap of 100 bp
* 14771 15475: contig of 705 bp in length
* 15476 15575: gap of 100 bp
* 15576 16289: contig of 714 bp in length
* 16290 16389: gap of 100 bp

16390 17130: contig of 741 bp in length
* 17131 17230: gap of 100 bp
* 17231 17963: contig of 733 bp in length
* 17964 18063: gap of 100 bp
* 18064 18768: contig of 705 bp in length
* 18769 18868: gap of 100 bp
* 18869 19597: contig of 729 bp in length
* 19598 19697: gap of 100 bp
* 19698 20447: contig of 750 bp in length
* 20448 20547: gap of 100 bp
* 20549 21283: contig of 736 bp in length
* 21284 21383: gap of 100 bp
* 21384 21421: contig of 738 bp in length
* 21422 22221: gap of 100 bp
* 22222 22943: contig of 722 bp in length
* 22944 23043: gap of 100 bp
* 23044 23755: contig of 712 bp in length
* 23756 23855: gap of 100 bp
* 23856 24596: contig of 741 bp in length
* 24597 24696: gap of 100 bp
* 24697 25409: contig of 713 bp in length
* 25410 25509: gap of 100 bp
* 25510 26208: contig of 699 bp in length
* 26209 26308: gap of 100 bp
* 26309 27001: contig of 693 bp in length
* 27002 27101: gap of 100 bp
* 27102 27831: contig of 730 bp in length
* 27832 27931: gap of 100 bp
* 27932 28661: contig of 730 bp in length
* 28662 28761: gap of 100 bp
* 28762 29500: contig of 739 bp in length
* 29501 29600: gap of 100 bp
* 29601 30333: contig of 733 bp in length
* 30334 30433: gap of 100 bp
* 30434 31135: contig of 702 bp in length
* 31136 31235: gap of 100 bp
* 31236 31945: contig of 710 bp in length
* 31946 32045: gap of 100 bp
* 32046 32757: contig of 712 bp in length
* 32758 32857: gap of 100 bp
* 32858 33581: contig of 724 bp in length
* 33582 33681: gap of 100 bp
* 33682 34361: contig of 680 bp in length
* 34362 34461: gap of 100 bp
* 34462 35166: contig of 705 bp in length
* 35167 35266: gap of 100 bp
* 35267 35993: contig of 727 bp in length
* 35994 36093: gap of 100 bp
* 36094 36829: contig of 736 bp in length
* 36830 36929: gap of 100 bp
* 36930 37662: contig of 733 bp in length
* 37663 37762: gap of 100 bp
* 37763 38501: contig of 739 bp in length
* 38502 38601: gap of 100 bp
* 38602 39342: contig of 741 bp in length
* 39343 39442: gap of 100 bp
* 39443 40161: contig of 719 bp in length
* 40162 40261: gap of 100 bp
* 40262 40982: contig of 721 bp in length
* 40983 41082: gap of 100 bp
* 41083 41816: contig of 734 bp in length
* 41817 41916: gap of 100 bp
* 41917 42623: contig of 707 bp in length
* 42624 42723: gap of 100 bp
* 42724 43429: contig of 706 bp in length
* 43430 43529: gap of 100 bp
* 43530 44262: contig of 733 bp in length
* 44263 44362: gap of 100 bp
* 44363 45097: contig of 735 bp in length
* 45098 45197: gap of 100 bp
* 45198 45926: contig of 729 bp in length
* 45927 46026: gap of 100 bp
* 46027 46744: contig of 718 bp in length

TITLE
JOURNAL
COMMENT

```

* 46745 46844: gap of 100 bp
* 47580: contig of 736 bp in length
* 47581 47680: gap of 100 bp
* 48418: contig of 738 bp in length
* 48519 48518: gap of 100 bp
* 49224: contig of 706 bp in length
* 49325 49324: gap of 100 bp
* 50043 50042: contig of 718 bp in length
* 50143 50142: gap of 100 bp
* 50877 50876: contig of 734 bp in length
* 51693 51692: gap of 100 bp
* 51793 51792: contig of 716 bp in length
* 52597 52596: gap of 100 bp
* 53307 53306: contig of 710 bp in length
* 53407 53406: gap of 100 bp
* 54111 54110: contig of 704 bp in length
* 54211 54210: gap of 100 bp
* 54949 54948: contig of 738 bp in length
* 55049 55048: gap of 100 bp
* 55777 55776: contig of 728 bp in length
* 55877 55876: gap of 100 bp
* 56620: contig of 744 bp in length

Query Match
Best Local Similarity 1.0%; Score 33; DB 12; Length 72149;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2008 GAATTTGGGCCACCATTTTGGATTGACGAGGT 2040
Db 1714 GAATTTGGGCCACCATTTTGGATTGACGAGGT 1746

RESULT 77
CR353491 1449 bp mRNA linear VRT 11-MAR-2004
LOCUS Gallus gallus finished cDNA, clone CHEST24919.
DEFINITION CR353491
ACCESSION CR353491
VERSION CR353491.1 GI:45424761
KEYWORDS Gallus gallus (chicken)
SOURCE Gallus gallus
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus
1 (bases 1 to 1449)
Boardman P.E., Bonfield, J.K., Brown, W.R.A., Carder, C., Chalk, S.E.,
Croning, M.D.R., Davies, R.M., Francie, M.D., Grafham, D.V.,
Hubbard, S.U., Humphray, S.J., Hunt, P.J., Maddison, M., McLaren, S.R.,
Niblett, D., O'Connell, P., Rogers, J., Scott, C.E., Taylor, R.G.,
Tickle, C., and Wilson, S.A.
Direct Submission
Submitted (09-MAR-2004) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: chickest@ms.umist.ac.uk
BBSRC/bundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
sequencing project.
This sequence is from the
BBSRC/bundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
from a library constructed by Elizabeth Bosch. cDNA was prepared
from RNA extracted from liver, normalised, and poly A-trimmed.
EcoRI-NotI cut cDNA was then ligated into the vector. Vector:
pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
coli DH10B.
Location/Qualifiers
1..1449
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHEST24919"
/clone_lib="CSEQCHN33"
/dev_stage="adult"
FEATURES
source

```

ORIGIN

```

Query Match
Best Local Similarity 1.0%; Score 32; DB 11; Length 1449;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2458 TTAATGCAGAGTCAGATATCTTCAGGTTCC 2489
Db 815 TTAATGCAGAGTCAGATATCTTCAGGTTCC 846

```

RESULT 78

```

AC163387 194188 bp DNA linear ROD 13-JUL-2005
LOCUS Mus musculus chromosome 1, clone RP24-181G14, complete sequence.
DEFINITION AC163387
ACCESSION AC163387
VERSION AC163387.2 GI:70778622
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 194188)
Birren, B., Nusbaum, C. and Lander, E.
Mus musculus chromosome 1, clone RP24-181G14
Unpublished
2 (bases 1 to 194188)

```

REFERENCE

```

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J.,
Choepe, Y., Collamore, A., Cook, A., Cooke, P., Corum, B.,
Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D.,
Galligan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kell, C., Landers, T.,
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, R., Mabbitt, R.,
McCarthy, M., McDonald, P., Major, J., Manning, J., Matthews, C.,
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
Zimmer, A. and Zody, M.
Direct Submission
Submitted (10-JUN-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 194188)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J.,
Choepe, Y., Collamore, A., Cook, A., Cooke, P., Corum, B.,
Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D.,
Galligan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kell, C., Landers, T.,
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, R., Mabbitt, R.,
McCarthy, M., McDonald, P., Major, J., Manning, J., Matthews, C.,
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
Zimmer, A. and Zody, M.
Direct Submission
Submitted (10-JUN-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 194188)

```

TITLE

JOURNAL

REFERENCE

AUTHORS

us-10-825-632-2.olig.rge

Thu Jun 22 09:04:25 2006

TITLE Direct Submission
JOURNAL Submitted (13-JUL-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Jul 13, 2005 this sequence version replaced gi:67460016. All repeats were identified using RepeatMasker.
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Broad Institute of MIT and Harvard
Center code: WTRR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: L32375
Center clone name: 181_G_14

FEATURES
source
1.194188
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clone="RP24-181G14"
/clone.lib="RPCI-24 Male Mouse BAC"
215..252
/rpt_family="A-rich"
repeat_region
3573..4002
/rpt_family="Lx7"
repeat_region
4729..4975
/rpt_family="Lx7"
repeat_region
5173..5207
/rpt_family="GA)n"
repeat_region
6288..6371
/rpt_family="L1_MM"
repeat_region
6619..6682
/rpt_family="TA)n"
repeat_region
7825..7870
/rpt_family="CT-rich"
repeat_region
8855..8925
/rpt_family="CAT)n"
repeat_region
complement(8933..9149)
/rpt_family="ORR1D"
repeat_region
complement(9634..9794)
/rpt_family="ORR1D"
repeat_region
9795..9912
/rpt_family="CA)n"
repeat_region
complement(9919..9973)
/rpt_family="L1M2"
repeat_region
complement(10172..10316)
/rpt_family="B1_MM"
repeat_region
10891..10955
/rpt_family="BC1_MM"
repeat_region
complement(11247..11357)
/rpt_family="Lx9"
repeat_region
complement(12029..12441)
/rpt_family="RLTR9B"
repeat_region
12748..12769
/rpt_family="AT-rich"
repeat_region
complement(12770..17468)
/rpt_family="L1_RN"
repeat_region
17477..18041
/rpt_family="RMER13B"
repeat_region
complement(17711..18496)
/rpt_family="RMER17B"
repeat_region
complement(18497..19986)
/rpt_family="L1VL2"
repeat_region
20054..20120
/rpt_family="L2"
repeat_region
23394..23524
/rpt_family="L2"
complement(23749..23955)

repeat_region
24591..24670
/rpt_family="L1F"
repeat_region
24665..25207
/rpt_family="L1"
repeat_region
25201..30672
/rpt_family="L1_MM"
repeat_region
30798..30833
/rpt_family="AT-rich"
complement(30965..31129)
/rpt_family="ORR1A2"
34254..34283
/rpt_family="AT-rich"
repeat_region
34720..34857
/rpt_family="RLTR1"
repeat_region
34906..35074
/rpt_family="RLTR1"
repeat_region
35074..35230
/rpt_family="RLTR1"
repeat_region
37290..37986
/rpt_family="Lx4"
repeat_region
37997..38021
/rpt_family="TAAA)n"
repeat_region
38217..38258
/rpt_family="(TG)n"
repeat_region
38258..38318
/rpt_family="(GA)n"
repeat_region
38555..39775
/rpt_family="Lx4"
repeat_region
39792..39892
/rpt_family="RLTR13A"
repeat_region
39893..40769
/rpt_family="RLTR13B"
repeat_region
40773..41089
/rpt_family="Lx4"
repeat_region
41101..41169
/rpt_family="CT-rich"
complement(41175..41247)
/rpt_family="RLTR11A"
repeat_region
41248..41284
/rpt_family="(CA)n"
complement(41285..41570)
/rpt_family="RLTR11A"
repeat_region
41581..42364
/rpt_family="L1VL2"
repeat_region
42364..42418
/rpt_family="L1"
repeat_region
42841..44071
/rpt_family="L1"
repeat_region
44072..44115
/rpt_family="(TAAA)n"
repeat_region
44115..46091
/rpt_family="L1"
complement(46092..46283)
/rpt_family="MT2C"
repeat_region
46876..47091
/rpt_family="B3"
repeat_region
47134..47324
/rpt_family="(TCTCCC)n"
repeat_region
48006..48153
/rpt_family="B1_MM"
repeat_region
48154..48208
/rpt_family="(CAAAA)n"
complement(48286..48496)
/rpt_family="RMER3"
49024..49100
/rpt_family="(TC)n"

Query Match 0.9%; Score 29; DB 6; Length 194188;
Best Local Similarity 100.0%; Pred.No. 0.00043;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1072 TATGAGAAAATGATGAATCTGAGGTGGA 1100


```

methods of use thereof
JOURNAL Patent: WO 0147944-A 971 05-JUL-2001;
FEATURES Curagen Corporation (US)
source Location/Qualifiers
1. .51
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Accession number CG43920127"

ORIGIN
Query Match 0.8%; Score 25; DB 2; Length 51;
Best Local Similarity 100.0%; Pred.No. 0.082;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2756 GACACAGCATTAAGAGTTCTCGAATC 2780
Db 51 GACACAGCATTAAGAGTTCTCGAATC 27

RESULT 82
LOCUS CQ572930 9154 bp DNA linear PAT 02-FEB-2004
DEFINITION Sequence 688 from Patent WO0171042.
ACCESSION CQ572930
VERSION CQ572930.1 GI:41637266
KEYWORDS
SOURCE Drosophila sp.
ORGANISM Drosophila sp.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
AUTHORS Venter, J.C., Adams, M., Li, P.W. and Myers, E.W.
TITLE Detection kits, such as nucleic acid arrays, for detecting the
expression of 10,000 or more Drosophila genes and uses thereof
JOURNAL Patent: WO 0171042-A 688 27-SEP-2001;
PE Corporation (NY) (US)
FEATURES
source
1. .9154
/organism="Drosophila sp."
/mol_type="unassigned DNA"
/db_xref="taxon:7242"

ORIGIN
Query Match 0.8%; Score 24; DB 2; Length 9154;
Best Local Similarity 100.0%; Pred.No. 0.36;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 625 CTATTAGAGAGAAAGAAACGCATT 648
Db 1771 CTATTAGAGAGAAAGAAACGCATT 1794

RESULT 83
LOCUS CQ608996 40388 bp DNA linear PAT 02-FEB-2004
DEFINITION Sequence 36754 from Patent WO0171042.
ACCESSION CQ608996
VERSION CQ608996.1 GI:41661310
KEYWORDS
SOURCE Drosophila sp.
ORGANISM Drosophila sp.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
AUTHORS Venter, J.C., Adams, M., Li, P.W. and Myers, E.W.
TITLE Detection kits, such as nucleic acid arrays, for detecting the
expression of 10,000 or more Drosophila genes and uses thereof
JOURNAL Patent: WO 0171042-A 36754 27-SEP-2001;
PE Corporation (NY) (US)
FEATURES
Location/Qualifiers

methods of use thereof
JOURNAL Patent: WO 0147944-A 971 05-JUL-2001;
FEATURES Curagen Corporation (US)
source Location/Qualifiers
1. .51
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Accession number CG43920127"

ORIGIN
Query Match 0.8%; Score 24; DB 2; Length 40388;
Best Local Similarity 100.0%; Pred.No. 0.37;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 625 CTATTAGAGAGAAAGAAACGCATT 648
Db 13418 CTATTAGAGAGAAAGAAACGCATT 13441

RESULT 84
LOCUS AC015386 79001 bp DNA linear HTG 16-NOV-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.
ACCESSION AC015386
VERSION AC015386.1 GI:6435949
KEYWORDS HTG: HTGS PHASE2.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 79001)
AUTHORS Adams, M. and Venter, J.C.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT * NOTE: This information was identified as CDM:10210324 by the submitter.
This sequence was identified on this sequence e-mail to fly@celera.com.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1. .79001
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

ORIGIN
Query Match 0.8%; Score 24; DB 12; Length 79001;
Best Local Similarity 100.0%; Pred.No. 0.37;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 625 CTATTAGAGAGAAAGAAACGCATT 648
Db 68006 CTATTAGAGAGAAAGAAACGCATT 68029

RESULT 85
LOCUS AC090436 102653 bp DNA linear PLN 01-OCT-2002
DEFINITION Chlamydomonas reinhardtii clone cr-3hi, complete sequence.
ACCESSION AC090436
VERSION AC090436.37 GI:23396228
KEYWORDS HTG.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 102653)
AUTHORS Jia, H., Lin, S., Wu, H., Dutcher, S. and Roe, B.A.
TITLE Chlamydomonas reinhardtii BAC Clone cr-3hi
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 102653)
AUTHORS Jia, H., Lin, S., Wu, H., Dutcher, S. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2001) Department Of Chemistry And Biochemistry,

```

REFERENCE
AUTHORS
TITLE
JOURNAL

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 102653)
Jia,H., Lin,S., Wu,H., Dutcher,S. and Roe,B.A.
Direct Submission
Submitted (05-SEP-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 102653)
Jia,H., Lin,S., Wu,H., Dutcher,S. and Roe,B.A.
Direct Submission
Submitted (05-SEP-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
5 (bases 1 to 102653)
Jia,H., Lin,S., Wu,H., Dutcher,S. and Roe,B.A.
Direct Submission
Submitted (01-OCT-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Oct 1, 2002 this sequence version replaced gi:22748433.
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

COMMENT

FEATURES
source
Location/Qualifiers
1..102653
/organism="Chlamydomonas reinhardtii"
/mol_type="genomic DNA"
/db_xref="taxon:3055"
/clone="cr-3h1"
/clone_lib="cr-3h1"

ORIGIN

Query Match 0.8%; Score 24; DB 4; Length 102653;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 GCCGCTGCTACTGCCCGCTGCT 49
|||||
Db 18109 GCCGCTGCTACTGCCCGCTGCT 18086
|||||

RESULT 86
AC090434/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC090434 127686 bp DNA linear PLN 17-SEP-2003
Chlamydomonas reinhardtii clone cr-1j6, complete sequence.
AC090434
AC090434.43 GI:34787433
HTG.
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
Chlamydomonadales; Chlamydomonadales; Chlamydomonas.

REFERENCE
AUTHORS
TITLE
JOURNAL

Wu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.
Chlamydomonas reinhardtii BAC Clone cr-1j6
Unpublished

REFERENCE
AUTHORS
TITLE
JOURNAL

Wu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.
Direct Submission
Submitted (21-FEB-2001) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 127686)
Wu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.
Direct Submission
Submitted (03-AUG-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 127686)
Wu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.

REFERENCE
AUTHORS
TITLE
JOURNAL

Wu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.
Direct Submission
Submitted (03-AUG-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 127686)
Wu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.

TITLE
JOURNAL

Direct Submission
Submitted (14-AUG-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
5 (bases 1 to 127686)
Wu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.
Direct Submission
Submitted (17-SEP-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Sep 17, 2003 this sequence version replaced gi:22218469.
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES
source

Location/Qualifiers
1..127686
/organism="Chlamydomonas reinhardtii"
/mol_type="genomic DNA"
/db_xref="taxon:3055"
/clone="cr-1j6"
/clone_lib="cr-1j6"
/note="This is one of two clone in the same well cr-1j6"

ORIGIN

Query Match 0.8%; Score 24; DB 4; Length 127686;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 GCCGCTGCTACTGCCCGCTGCT 49
|||||
Db 102737 GCCGCTGCTACTGCCCGCTGCT 102714
|||||

RESULT 87
LOCUS

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC010110 172372 bp DNA linear INV 24-JUL-2002
Drosophila melanogaster 3L BAC RP98-9N2 (Rowell Park Cancer
Institute Drosophila BAC Library) complete sequence.
AC010110
AC010110.7 GI:21553183
HTG.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS

Muzny,D., Scherer,S., Adams,M.D., Holt,R.A., Evans,C.A.,
Gocayne,J.D., Tabor,P., Williamson,A., Homsai,F.H., Chen,C.,
Dugan-Rocha,S.D., Sodergren,E.S., Hodgson,A.H., Chen,B.C.,
Ayele,M., Scott,G.S., Worley,K.W., Anamstides,P.G., Brandon,R.C.,
Rogers,Y., An,H., Baldwin,D., Beeson,K.Y., Brown,M., Bunay,C.,
Busam,D.A., Center,A., Chen,G., Chen,Z., Clerk-Blankenburg,K.,
Davenport,L.B., Dietz,S.M., Ding,Y., Dodson,K., Doup,L.E.,
Draper,H., Emery-Cohen,A., Fierster,J.S., Garg,N.D.S., Houck,J.,
Hostin,D., Howland,T.J., Hume,J., Ibegwam,C., Jaisi,M., Kovar,C.,
Liu,W., Mattei,B., McIntosh,I.C., Morgan,M., Moy,M., Murphy,B.,
Nelson,K.A., Nkasesa,Y., Nguyen,N., Perez,L., Pittman,G.S., Puri,V.,
Xiang,J., Shen,H., Strong,R., Tector,C., Wang,Q., Williams,S.M.,
Wheeler,D., Zaveri,J.S., Zhou,J., Zorrilla,S., Smith,H.O.,
Direct Submission
Unpublished

TITLE
JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 172372)
Worley,K.C., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooks,S.L., Amarutunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Blinag,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brivea,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Cartier,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,

Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulsged, H., Lozano, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G. and Gibbs, R.

Direct Submission
Unpublished
3 (bases 1 to 172372)
Worley, K.C.

Direct Submission
Submitted (13-SEP-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 172372)
Worley, K.C., Adams, C., Adio-Odola, B., Ali-oman, F.R., Allen, C., Albrook, S.L., Amarutunge, H.C., Are, J.R., Banks, F., Barbara, J., Benton, J., Bimaga, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Bowie, S., Brivaga, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulsged, H., Lozano, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H.,

Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G. and Gibbs, R.

Direct Submission
Submitted (31-AUG-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

5 (bases 1 to 172372)
BCM-HGSC.

Direct Submission
Submitted (25-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

6 (bases 1 to 172372)
BCM-HGSC.

Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Jun 25, 2002 this sequence version replaced gi:15383754.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht>

FEATURES	Location/Qualifiers
source	1..172372 /organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /chromosome="3L" /clone="RP98-9N2" 101164..101604 /function="pcr product sequence only"
misc_feature	101451..101573 /function="single clone coverage"
misc_feature	
ORIGIN	

```

Query Match      0.8%; Score 24; DB 13; Length 172372;
Best Local Similarity 100.0%; Pred. NO. 0.38;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 625 CTATTACAGAGAAAGAAACGCATT 648
Db 88694 CTATTACAGAGAAAGAAACGCATT 8917

RESULT 88
AC127924
LOCUS      298235 bp DNA linear HTG 19-NOV-2002
DEFINITION Rattus norvegicus clone CH230-23318, *** SEQUENCING IN PROGRESS
AC127924.3 GI:25075391
HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.
AC127924
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
1 (bases 1 to 298235)
Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Amin,A., Anguiano,D.,
Allen,C., Allen,H., Alsbrooks,S., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaraike,D., Barber,M., Barneshead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,R.,
Cardenas,V., Carter,K., Cavazos,J., Ceasar,H., Chen,R.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyne,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,X., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finlay,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Georgievski,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,M., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hoques,M.,
Hollins,B., Howells,S., Huliyil,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshewa,L., Louised,H., Lozada,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,X.,
Nwaokeme,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plummer,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Shen,H.,
Sander,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajd,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Soza,D.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,
Taylor,T., Thomas,R., Thomas,S., Tingey,A., Trejos,Z., Umanik,K.,
Valas,R., Vexa,V., Villanasa,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Zimm,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission
Unpublished
2 (bases 1 to 298235)
Worley,K.C.

Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 298235)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:23915349.

The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: KAAS

Center clone name: CH230-23318

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 245203 bases at least Q40

Consensus quality: 249375 bases at least Q30

Consensus quality: 252157 bases at least Q20

Estimated insert size: 248526; sum-of-contigs estimation

Quality coverage: 5x in Q20 bases, sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 15 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N. But the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 9520: contig of 9520 bp in length

9521 9520: gap of unknown length

9522 24965: contig of 15345 bp in length

24966 25065: gap of unknown length

25066 43623: contig of 18558 bp in length

43624 43723: gap of unknown length

43724 58040: contig of 14317 bp in length

58041 58140: gap of unknown length

58141 276319: contig of 218079 bp in length

276320 276319: gap of unknown length

276321 281371: contig of 4952 bp in length

281372 282614: contig of 1243 bp in length

282615 282714: gap of unknown length

282715 283928: contig of 1214 bp in length

283929 284028: gap of unknown length

284029 285169: contig of 1141 bp in length

285170 285269: gap of unknown length

285270 286516: contig of 1247 bp in length

286517 286516: gap of unknown length

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES	source	1. 286617 287814: contig of 1198 bp in length	AB003475.4 GI:55380495	VERSION
		287815 289938: gap of unknown length	Drosophila melanogaster (fruit fly)	KEYWORDS
misc_feature		287915 289938: contig of 1024 bp in length	Drosophila melanogaster	SOURCE
		289939 290451: gap of unknown length	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	ORGANISM
misc_feature		290452 290551: gap of unknown length	1 (bases 1 to 301691)	REFERENCE
		290552 292527: contig of 1976 bp in length	Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Chen,L.X., Sutton,G.G., Wortman,J.R., Vandeil,M.D., Zhang,Q., Henderson,S.N., Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Heit,G., Nelson,C.R., Gabor,G.L., Abril,J.F., Agbayani,A., An,H.J., Andrews-Frankoch,C., Baldwin,D., Ballaw,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferrier,S., Fleischmann,W., Fosler,C., Gabriellian,A.E., Garg,N.S., Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wei,W.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z., Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nuskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S., Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamoe,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,E., Svirskaas,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A., Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T., Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F., Zaveri,J.S., Zhan,M., Zhou,X., Zhu,S., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.	THE genome sequence of Drosophila melanogaster
misc_feature		281272 281371 /estimated_length=unknown	Science 287 (5461), 2185-2195 (2000)	TITLE
		282615 282714 /estimated_length=unknown	10731132	JOURNAL
gap		283929 284028 /estimated_length=unknown	2 (bases 1 to 301691)	PUBMED
		285170 285269 /estimated_length=unknown	Celniker,S.E., Wheeler,D.A., Krommiller,B., Carlson,J.W., Halpern,A., Patel,S., Adams,M., Champe,M., Dugan,S.P., Frise,E., Hodgson,A., George,R.A., Hoskins,R.A., Laverty,T., Muzny,D.M., Nelson,C.R., Pacleb,J.M., Park,S., Pfeiffer,B.D., Richards,S., Sodergren,E.J., Svirskaas,R., Tabor,P.E., Wan,K., Stapleton,M., Sutton,G.G., Venter,C., Weinstock,G., Scherer,S.E., Myers,E.W., Gibbs,R.A. and Rubin,G.M.	REFERENCES
gap		286517 286616 /estimated_length=unknown	Finishing a whole-genome shotgun: release 3 of the Drosophila melanogaster euchromatic genome sequence	TITLE
		287815 287914 /estimated_length=unknown	Genome Biol. 3 (12), RESEARCH0079 (2002)	JOURNAL
gap		288939 289038 /estimated_length=unknown	12537568	PUBMED
		290452 290551 /estimated_length=unknown	3 (bases 1 to 301691)	REFERENCE
gap		292528 292627 /estimated_length=unknown	Misra,S., Crosby,M.A., Mungall,C.J., Matthews,B.B., Campbell,K.S., Hradecky,P., Huang,Y., Kaminker,J.S., Millburn,G.H., Prochnik,S.E., Smith,C.D., Tupy,J.L., Whitfield,E.J., Bayraktaroglu,L., Berman,B.P., Bettencourt,B.R., Celniker,S.E., de Grey,A.D., Drysdale,R.A., Harris,N.L., Richter,J., Russo,S., Schroeder,A.J., Shu,S.Q., Stapleton,M., Yamada,C., Ashburner,M., Gelbart,W.M., Rubin,G.M. and Lewis,S.E.	AUTHORS
			Annotation of the Drosophila melanogaster euchromatic genome: a systematic review	TITLE
ORIGIN			Genome Biol. 3 (12), RESEARCH0083 (2002)	JOURNAL
			12537572	PUBMED
Query Match			4 (bases 1 to 301691)	REFERENCE
Best Local Similarity		0.87; Score 24; DB 12; Length 298235;		
		100.0%; Pred. No. 0.39;		
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY		26 GCCGCTGCTACTGCGCGCTGCT 49		
DB		221166 GCCGCTGCTACTGCGCGCTGCT 221189		
RESULT 89				
AE003475		301691 bp DNA linear INV 13-JAN-2006		
LOCUS				
DEFINITION		Drosophila melanogaster chromosome 3L, section 9 of 83 of the complete sequence.		
ACCESSION		AE003475 AE002584 AF014296		

AUTHORS	Kaminker, J.S., Bergman, C.M., Kronmiller, B., Carlson, J., Svirekas, R., Patel, S., Frise, E., Wheeler, D.A., Lewis, S.E., Rubin, G.M., Ashburner, M. and Celnikier, S.	
	The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective	
JOURNAL	Genome Biol.	3 (12), RESEARCH0084 (2002)
PUBMED	12537573	
REFERENCE	5 (bases 1 to 301691)	
AUTHORS	Celnikier, S., Carlson, J., Wan, K., Pfeiffer, B., Frise, E., George, R., Hoskins, R., Stapleton, J., Pacleb, J., Park, S., Svirekas, R., Smith, E., Yu, C. and Rubin, G.	
	Berkeley Drosophila Genome Project	
CONSRSTM	Drosophila melanogaster release 4 sequence	
JOURNAL	Unpublished	
REFERENCE	6 (bases 1 to 301691)	
AUTHORS	Adams, M.D., Celnikier, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.	
	Direct Submission	
JOURNAL	Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA	
REFERENCE	7 (bases 1 to 301691)	
AUTHORS	FlyBase	
	Direct Submission	
JOURNAL	Submitted (05-JAN-2006) FlyBase, Harvard University, Biological Laboratories, 16 Divinity Ave, Cambridge, MA 02138, USA	
COMMENT	On Nov 4, 2004 this sequence version replaced gi:23092840.	
FEATURES	Location/Qualifiers	
	1..301691	
Source	/organism="Drosophila melanogaster"	
	/mol_type="genomic DNA"	
Gene	/db_xref="taxon:7227"	
	/chromosome="3L"	
mRNA	267..>1262	
	/gene="CG33231"	
CDS	/locus_tag="Dmel_CG33231"	
	/old_locus_tag="CG33231"	
Gene	/map="62E3-62E3"	
	/db_xref="FLYBASE:FBgn0053231"	
mRNA	join(<267..688,1139..>1262)	
	/gene="CG33231"	
CDS	/locus_tag="Dmel_CG33231"	
	/old_locus_tag="CG33231"	
Gene	/product="CG33231-RA"	
	/db_xref="FLYBASE:FBgn0053231"	
mRNA	join(267..688,1139..>1262)	
	/gene="CG33231"	
CDS	/locus_tag="Dmel_CG33231"	
	/old_locus_tag="CG33231"	
Gene	/note="CG33231 gene product from transcript CG33231-RA"	
	/codon_start=1	
mRNA	/product="CG33231-PA"	
	/protein_id="AA564938.1"	
CDS	/db_xref="GI:45445784"	
	/translation="MCGRCFIMVLDLNLKRLKRSHTATVSNLRKATSADATRAA	
Gene	SDEVNPNQRLAQLIKNSVARRRRQQQQQQQQQPDGGHNSIRIRPSNKANSSSS	
	QCCSCSNRPVYTSFSPDIL"	
mRNA	2014..34339	
	/gene="CG33232"	
CDS	/locus_tag="Dmel_CG33232"	
	/old_locus_tag="CG33232"	
Gene	/map="62E3-62E4"	
	/db_xref="FLYBASE:FBgn0053347"	
mRNA	join(2014..3380,19675..19948,20184..20339,30122..30263,30330..30840,30912..31069,31129..33147,33210..33457,33517..34339)	
	/gene="CG33232"	
Gene	/locus_tag="Dmel_CG33232"	
	/old_locus_tag="CG33232"	
mRNA	/product="CG33232-RD, transcript variant D"	
	/db_xref="FLYBASE:FBgn0053347"	
Gene	complement(<9878..>10705)	
	/gene="CG16762"	
mRNA	/locus_tag="Dmel_CG16762"	
	/old_locus_tag="CG16762"	
CDS	/map="62E3-62E3"	
	/db_xref="FLYBASE:FBgn0035343"	
Gene	complement(join(<9878..10228,10292..>10705))	
	/gene="CG16762"	
mRNA	/locus_tag="Dmel_CG16762"	
	/old_locus_tag="CG16762"	
CDS	/product="CG16762-PA"	
	/protein_id="AAF47648.1"	
Gene	/db_xref="GI:7292239"	
	/translation="MCPMSVKILASLAMIALLVLSPAEGISRLPPVVGLPOTDGL	
mRNA	TRLFNLSRVTDTPFASVACFGYIGESNLIAELYSANVTKYNAADSRKIGIDADFL	
	ATRTTILSSRVCSBLRACNELNTLTFSQCHNVGNNTVSTYSISGNASASVL	
CDS	EERYVVDLRHGQCCQARAHVYESTARNYVLOACLGRKPKMPKPTSTSTT	
	TTTITTTTEAPPTPTTTEAPLNVEDQFKQLNLLN"	
Gene	complement(11253..13604)	
	/gene="Cyp4d20"	
mRNA	/locus_tag="Dmel_CG16761"	
	/old_locus_tag="CG16761"	
CDS	/map="62E3-62E3"	
	/db_xref="FLYBASE:FBgn0035344"	
Gene	complement(join(11253..12229,12296..12492,12555..12624,12801..13187,13257..13604))	
	/gene="Cyp4d20"	
mRNA	/locus_tag="Dmel_CG16761"	
	/old_locus_tag="CG16761"	
CDS	/product="CG16761-RA"	
	/db_xref="FLYBASE:FBgn0035344"	
Gene	complement(join(11514..12229,12296..12492,12555..12624,12801..13187,13257..13419))	
	/gene="Cyp4d20"	
mRNA	/locus_tag="Dmel_CG16761"	
	/old_locus_tag="CG16761"	
CDS	/note="CG16761 gene product from transcript CG16761-RA"	
	/codon_start=1	
Gene	/product="CG16761-PA"	
	/protein_id="RAP47649.1"	
mRNA	/db_xref="GI:7292240"	
	/db_xref="FLYBASE:FBgn0035344"	
CDS	Query Match 0.87; Score 24; DB 13; Length 301691;	
	Best Local Similarity 100.04; Pred.No. 0.39;	
Gene	Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
	QY 625 CTATTAAAGAGAAAGAAACGCATT 648	
mRNA	Db 275631 CTATTAAAGAGAAAGAAACGCATT 275654	
	RESULT 90	
Gene	CS196104	
	LOCUS	
mRNA	Sequence 3 from Patent WO2005106021.	
	Accession CS196104	
Gene	Version CS196104.1	
	GI:83410069	
mRNA	synthetic construct	
	other sequences; artificial sequences.	
Gene	REFERENCE	
	AUTHORS	
mRNA	Goiz, S.	
	Patent: WO 2005106021-A 3 10-NOV-2005;	

Query Match 0.7%; Score 23; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.2; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 213 CATGGCAGCAGCAATGGAAACAG 235
Db 7 CATGGCAGCAGCAATGGAAACAG 29

RESULT 93
BV055847 632 bp DNA linear STS 31-MAY-2003
LOCUS BV055847.1 GI:31171642
DEFINITION tagged site.
ACCESSION BV055847
VERSION BV055847.1
KEYWORDS STS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 632)
Wade, C.M., Kulbokas, E.J. III, Kirby, A.W., Zody, M.C., Mullikin, J.C.,
Lander, E.S., Lindblad-Toh, K. and Daly, M.J.
The mosaic structure of variation in the laboratory mouse genome
Nature 420 (6915), 574-578 (2002)
12466852

CONTACT: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 632
Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads
were generated from 129S1/SvTmJ, C3H/HeJ, and BALB/cByJ. The WGS
reads were placed uniquely on the MGSV3 C57BL/6J assembly and SNP
detection was carried out by SSAHA-SNP. 225,000 reads were
annotated
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
and the strain from which the particular read came. The validation
rate for these SNPs was estimated at approximately 98%.

FEATURES
source
1..632
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="CZECHII/Ei"
/db_xref="taxon:10090"
/map="+" 1 21-583 7608303-7608864"
/clone_lib="CZECHII/Ei"
<1..>632

STS
ORIGIN
Query Match 0.7%; Score 23; DB 7; Length 632;
Best Local Similarity 100.0%; Pred. No. 1.3; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1477 ATCTATGAGAAACACACAT 1499
Db 417 ATCTATGAGAAACACACAT 439

RESULT 94
BV311981 635 bp DNA linear STS 26-JAN-2005
LOCUS BV311981
DEFINITION S236P628FH12.TO Alaskan Malamute Canis familiaris STS genomic,
sequence tagged site.

Query Match 0.7%; Score 23; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.2; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 887 TCGTAACACAGAGAGAAAGGAGA 909
Db 1 TCGTAACACAGAGAGAAAGGAGA 23

RESULT 91
AR631300 29 bp DNA linear PAT 14-FEB-2005
LOCUS AR631300
DEFINITION Sequence 45 from patent US 6844180.
ACCESSION AR631300
VERSION AR631300.1
KEYWORDS GI:59770958
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 29)
Oi, S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 45 18-JAN-2005;
Ferring BV;
NLX;

FEATURES
source
1..29
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 0.7%; Score 23; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.2; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 213 CATGGCAGCAGCAATGGAAACAG 235
Db 7 CATGGCAGCAGCAATGGAAACAG 29

RESULT 92
AX608768 29 bp DNA linear PAT 17-FEB-2003
LOCUS AX608768
DEFINITION Sequence 45 from Patent WO0231134.
ACCESSION AX608768
VERSION AX608768.1
KEYWORDS GI:28404320
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.
TITLE Novel serine protease genes related to dppiv
JOURNAL Patent: WO 0231134-A 45 18-APR-2002;
Ferring BV (NL)

FEATURES
source
1..29
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

ACCESSION BV311981
 VERSION BV311981.1 GI:57510473
 STS
 SOURCE Canis familiaris (dog)
 ORGANISM Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 Canis
 1 (bases 1 to 635)
 The genome sequence of Canis familiaris
 Lindblad-Toh, K.
 Unpublished (2004)
 Contact: Kerstin Lindblad-Toh
 Whitehead Institute for Biomedical Research, Center for Genome
 Research
 320 Charles Street, Cambridge, MA 02141, USA
 Tel: 6172521477
 Fax: 6172580903
 Email: kersli@genome.wi.mit.edu
 Primer A: No sequence submitted
 Primer B: No sequence submitted
 STS size: 635
 Protocol:
 WGS-discovery (WGS):
 Paired-end and low-coverage whole genome shotgun reads were generated
 from 9 breeds
 (German Shepherd, Rottweiler, Bedlington Terrier, Beagle, Labrador
 Retriever, English
 Shepherd, Italian Greyhound, Alaskan Malamute and the Portuguese
 Water Dog -100,000 each)
 and five other canids (Chinese, Alaskan, Indian and Spanish Gray
 Wolf as well as the
 Californian Coyote).
 The WGS reads were placed uniquely on the CanFam1.0 boxer assembly
 and SNP detection was
 carried out by SSAHA-SNP. 863872 reads were annotated as STSs and
 485941 SNPs were
 annotated with alleles from the boxer and the breed or canid from
 which the particular
 read came. The validation rate for these SNPs was estimated at
 approximately 98%.
 WGA-discovery (WGA) of Boxer/Poodle SNPs:
 A second set of SNPs was generated using a similar methodology
 except that the contigs
 from the 1.5x poodle assembly (Kirkness 2003) were used instead of
 WGS reads. Since this
 phred 46 lacked base quality scores, arbitrary quality scores of
 before the poodle sequence was placed uniquely on the CanFam1.0
 boxer assembly and SNP
 detection was carried out by SSAHA-SNP. 1637780 SNPs were annotated
 with alleles from the
 boxer and the poodle. The validation rate for these SNPs was
 estimated at approximately TBD%.
 Internal-WGA-discovery (I-WGA):
 A third set of SNPs were discovered by comparing reads in the WGA
 assembly. SNPs were
 defined as mismatch positions that had a base quality of >= 30 on
 both reads in a region
 that aligned without gaps, and with at most one additional mismatch
 in the ten flanking
 bases. For each allele, at least one additional read had to confirm
 it. 731476 SNPs were
 annotated with alleles between the two boxer alleles. The
 validation rate for these SNPs
 was estimated at approximately TBD%.
 Location/Qualifiers
 1..635
 /organism="Canis familiaris"
 /mol_type="genomic DNA"
 /strain="AlaskanMalamute"
 /db_xref="taxon:9615"

ORIGIN
 Query Match 0.7%; Score 23; DB 7; Length 635;
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3098 ATAAGANTTACTTAAAAA 3120
 Db 595 ATAGANTTACTTAAAAA 617
 RESULT 95
 CS211305/c
 LOCUS CS211305 1275 bp DNA linear PAT 15-DEC-2005
 DEFINITION Sequence 624 from Patent WO2005111246.
 ACCESSION CS211305
 VERSION CS211305.1 GI:83691295
 KEYWORDS Mesocricetus auratus (golden hamster)
 ORGANISM Mesocricetus auratus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Cricetidae; Cricetinae; Mesocricetus.
 1
 Melville, M.W.
 Oligonucleotide arrays to monitor gene expression and methods for
 making and using same
 Patent: WO 2005111246-A 624 24-NOV-2005;
 Wyeth (US)
 FEATURES
 source
 1..1275
 /organism="Mesocricetus auratus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:10036"
 misc_feature 389..407
 /note="n is a, c, g, or t"
 misc_feature 457..476
 /note="n is a, c, g, or t"
 ORIGIN
 Query Match 0.7%; Score 23; DB 2; Length 1275;
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 26 GCCGCTGCTACTGCCGCGCTGC 48
 Db 501 GCCGCTGCTACTGCCGCGCTGC 479
 RESULT 96
 AB056577/c
 LOCUS AB056577 1275 bp mRNA linear ROD 25-SEP-2004
 DEFINITION Mesocricetus auratus mRNA for homeobox B3, complete cds.
 ACCESSION AB056577
 VERSION AB056577.1 GI:13359292
 KEYWORDS Mesocricetus auratus (golden hamster)
 ORGANISM Mesocricetus auratus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Cricetidae; Cricetinae; Mesocricetus.
 1
 Nakamura, N., Yoshimi, T. and Miura, T.
 Increased gene expression of lung marker proteins in the homeobox
 B3-overexpressed fetal lung cell line M3E3/C3
 Cell Growth Differ. 13 (4), 195-203 (2002)
 11971819
 JOURNAL
 PUBMED
 REFERENCE
 2 (bases 1 to 1275)
 Nakamura, N. and Miura, T.
 Direct Submission

/map="+ 6 22-578 54642973-54643529"
 /clone_lib="AlaskanMalamute"
 <1..>635
 STS
 ORIGIN
 Query Match 0.7%; Score 23; DB 7; Length 635;
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3098 ATAAGANTTACTTAAAAA 3120
 Db 595 ATAGANTTACTTAAAAA 617
 RESULT 95
 CS211305/c
 LOCUS CS211305 1275 bp DNA linear PAT 15-DEC-2005
 DEFINITION Sequence 624 from Patent WO2005111246.
 ACCESSION CS211305
 VERSION CS211305.1 GI:83691295
 KEYWORDS Mesocricetus auratus (golden hamster)
 ORGANISM Mesocricetus auratus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Cricetidae; Cricetinae; Mesocricetus.
 1
 Melville, M.W.
 Oligonucleotide arrays to monitor gene expression and methods for
 making and using same
 Patent: WO 2005111246-A 624 24-NOV-2005;
 Wyeth (US)
 FEATURES
 source
 1..1275
 /organism="Mesocricetus auratus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:10036"
 misc_feature 389..407
 /note="n is a, c, g, or t"
 misc_feature 457..476
 /note="n is a, c, g, or t"
 ORIGIN
 Query Match 0.7%; Score 23; DB 2; Length 1275;
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 26 GCCGCTGCTACTGCCGCGCTGC 48
 Db 501 GCCGCTGCTACTGCCGCGCTGC 479
 RESULT 96
 AB056577/c
 LOCUS AB056577 1275 bp mRNA linear ROD 25-SEP-2004
 DEFINITION Mesocricetus auratus mRNA for homeobox B3, complete cds.
 ACCESSION AB056577
 VERSION AB056577.1 GI:13359292
 KEYWORDS Mesocricetus auratus (golden hamster)
 ORGANISM Mesocricetus auratus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Cricetidae; Cricetinae; Mesocricetus.
 1
 Nakamura, N., Yoshimi, T. and Miura, T.
 Increased gene expression of lung marker proteins in the homeobox
 B3-overexpressed fetal lung cell line M3E3/C3
 Cell Growth Differ. 13 (4), 195-203 (2002)
 11971819
 JOURNAL
 PUBMED
 REFERENCE
 2 (bases 1 to 1275)
 Nakamura, N. and Miura, T.
 Direct Submission

JOURNAL Submitted (01-MAR-2001) Nobuatsu Nakamura, Tokyo University of Pharmacy and Life Science, Department of Environmental Molecular Physiology; Horinouchi 1432-1, Hachiohji, Tokyo 192-0392, Japan (E-mail: s94804@educ.its.toyaku.ac.jp, URL: http://www.its.toyaku.ac.jp/Life-Science/lemp-4/, Tel: 81-426-76-5053, Fax: 81-426-76-6811)

FEATURES

source

1..1275
/location/Qualifiers
/organism="Mesocricetus auratus"
/mol_type="mRNA"
/db_xref="taxon:10036"
/cell_line="M3E3/C3"

gene

1..1275
/gene="hox b3"

CDS

1..1275
/gene="hox b3"
/codon_start=1
/product="homeobox B3"
/protein_id="BAB33375.1"
/db_xref="GI:13359293"
/translation="MQKATYYNTAAALFGGYSSYPGNGFYDGPFPQPPQAATHLE GDYORASCSLQSNAAAPAKSKELNGSCMRPLAPELPAPGSPPSAAPTSTTNS SNGGSGSGKPKCGPGSGNSTLTQIFPMKESROTSLKNSSPGTGCGGGGGGG SSGSSGGSGGGGGKSPGSAASKRARTAYTSAQLVELEKEFHFRYLCPRRIVE MANLLLSERQIKIWFQNRMKYKQOKAKGLASSGGSPAGSPQPMQSTAGFMNA LHSMTSPSDSPSPAPATKGHONAYALPSNVOPPLKGCAGAPQKCPPTPASEYEPHVLQA NGGAYGTPTMQGSFVYGGGVADPLPPAGPSLYGLNHLSHHPSGLNDYNGAAPMPG NQHRGPCDHPFTITLSSHAPPQGRQIAPN"

misc_feature

385..402
/gene="hox b3"
/note="hexa peptide"

misc_feature

556..738
/gene="hox b3"
/note="homeodomain"

ORIGIN

Query Match 0.7%; Score 23; DB 6; Length 1275;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 GCCGCTGCTACTGCCGCGCTGC 48
|||||

Db 501 GCCGCTGCTACTGCCGCGCTGC 479
|||||

RESULT 97

AC149896 74579 bp DNA linear VRT 02-SEP-2005
LOCUS Xenopus tropicalis clone ISB-27807, complete sequence.
DEFINITION AC149896
ACCESSION AC149896.2 GI:74096560
VERSION HTG.
KEYWORDS
SOURCE Xenopus tropicalis (Silurana tropicalis)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Silurana.

REFERENCE

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL
REFERENCE 2 (bases 1 to 74579)
AUTHORS DOE Joint Genome Institute
TITLE Direct Submission
JOURNAL Submitted (24-JUN-2004) Production Genomics Facility, DOE Joint Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA

REFERENCE

AUTHORS Stanford Human Genome Center.
AUTHORS DOE Joint Genome Institute
TITLE Direct Submission
JOURNAL Submitted (02-SEP-2005) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT On Sep 2, 2005 this sequence version replaced gi:49170153. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center www.shgc.stanford.edu
Quality: Phrap Quality >=40 98.8% of Sequence;
Estimated Total Number of Errors is 0.5.

FEATURES

source

1..74579
/location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB-27807"

ORIGIN

Query Match 0.7%; Score 23; DB 11; Length 74579;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2526 TGATACAGATACACGGAACGTT 2548
|||||

Db 55320 TGATACAGATACACGGAACGTT 55298
|||||

RESULT 98

AC149892 98359 bp DNA linear HTG 24-JUN-2004
LOCUS Xenopus tropicalis clone ISB-242D2, WORKING DRAFT SEQUENCE, 8
DEFINITION AC149892
ACCESSION AC149892
VERSION AC149892.1 GI:49170149
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Xenopus tropicalis (Silurana tropicalis)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Silurana.

REFERENCE

AUTHORS DOE Joint Genome Institute.
TITLE Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 98359)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (24-JUN-2004) Production Genomics Facility, DOE Joint Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA

COMMENT

-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 2865953
Center clone name: ISB-242D2

Summary Statistics
Consensus quality: 95112 bases at least Q40
Consensus quality: 96157 bases at least Q30
Consensus quality: 96888 bases at least Q20
Estimated insert size: 104000; agarose-fp estimation
Estimated insert size: 97659; sum-of-contigs estimation
Quality coverage: 15.53 in Q20 bases; agarose-fp estimation
Quality coverage: 16.54 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

* as soon as it is available and the accession number will
* be preserved.

1 2277: contig of 2277 bp in length
2278: gap of unknown length
2378 2377: gap of unknown length
5120: contig of 2743 bp in length
5121 5220: gap of unknown length
5221 8464: contig of 3244 bp in length
8465 8564: gap of unknown length
8565 19956: contig of 11392 bp in length
19957 20056: gap of unknown length
20057 30478: contig of 10422 bp in length
30479 30578: gap of unknown length
30579 41701: contig of 11122 bp in length
41701 52926: gap of unknown length
52926 53025: contig of 11125 bp in length
53025 53026: gap of unknown length
53026 98359: contig of 45334 bp in length.

FEATURES

Location/Qualifiers
1..98359
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB-242D2"
2278..2377
/estimated_length=unknown
5121..5220
/estimated_length=unknown
8465..8564
/estimated_length=unknown
19957..20056
/estimated_length=unknown
30479..30578
/estimated_length=unknown
41701..41800
/estimated_length=unknown
52926..53025
/estimated_length=unknown

ORIGIN

Query Match 0.7%; Score 23; DB 12; Length 98359;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2526 TGATACAGGATACACGGAACGTT 2548
Db 55017 TGATACAGGATACACGGAACGTT 55039
|||||

RESULT 99
AC161463 213261 bp DNA linear VRT 16-DEC-2005
LOCUS Gallus gallus BAC clone CH261-127J9 from chromosome ul, complete
sequence.

AC161463
AC161463.3 GI:83699520
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus

1 (bases 1 to 213261)
Delgado B, Kozlowski A, and Lewis S.
The sequence of Gallus BAC clone CH261-127J9
Unpublished (2001)
2 (bases 1 to 213261)
Wilson R.K.

Direct Submission
Submitted (14-MAY-2005) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 213261)
Wilson R.K.
Direct Submission

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE

JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (29-OCT-2005) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 213261)
Wilson R.K.
Direct Submission
Submitted (16-DEC-2005) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA
On Dec 16 2005 this sequence version replaced gi:78272069.
-----Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUOSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
-----Summary Statistics
Center project name: J_AA127J09

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

SOURCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

Location/Qualifiers
1..213261
/organism="Gallus gallus"
/mol_type="genomic DNA"
/db_xref="taxon:9031"
/chromosome="ul"
/clone="CH261-127J9"
/clone_lib="CHORI261"

ORIGIN

Query Match 0.7%; Score 23; DB 11; Length 213261;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1477 ATCTATGAAGAAACAACAGACAT 1499
Db 25695 ATCTATGAAGAAACAACAGACAT 25717
|||||

RESULT 100
AR631303 22 bp DNA linear PAT 14-FEB-2005
LOCUS AR631303
DEFINITION Sequence 48 from patent US 6844180.
ACCESSION AR631303
VERSION AR631303.1 GI:59770964
KEYWORDS
SOURCE
ORGANISM

Unknown.
Unknown.
Unclassified.

REFERENCE
1 (bases 1 to 22)
Oi S., Akinsanya K.O., Riviere P.J.M. and Junien J.-L.
Serine protease genes related to DpPrV
Patent: US 6844180-A 48 18-JAN-2005;
Ferring BV,;
NLX,

FEATURES

Location/Qualifiers
1..22
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 0.7%; Score 22; DB 2; Length 22;

Thu Jun 22 09:04:25 2006

Best Local Similarity 100.0%; Pred. No. 4.7; Mismatches 0; Indels 0; Gaps 0;

QY 1880 CACATTCCTGTCATCAGTCA 1901
 |||
 DB 1 CACATTCCTGTCATCAGTCA 22

RESULT 101
 AR631304/c
 LOCUS AR631304 22 bp DNA linear PAT 14-FEB-2005
 DEFINITION Sequence 49 from patent US 6844180.
 ACCESSION AR631304
 VERSION AR631304.1 GI:59770966
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 22)
 AUTHORS Oi, S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.
 TITLE Serine protease genes related to DPPIV
 JOURNAL Patent: US 6844180-A 49 18-JAN-2005;
 Ferring BV;;
 FEATURES Location/Qualifiers
 source 1..22
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
 Query Match 0.7%; Score 22; DB 2; Length 22;
 Best Local Similarity 100.0%; Pred. No. 4.7; Mismatches 0; Indels 0; Gaps 0;
 Matches 22; Conservative 0;

QY 1972 TCAAGTCCTGAAGATGACCCAA 1993
 |||
 DB 22 TCAAGTCCTGAAGATGACCCAA 1

RESULT 102
 AX608771
 LOCUS AX608771 22 bp DNA linear PAT 17-FEB-2003
 DEFINITION Sequence 48 from Patent WO0231134.
 ACCESSION AX608771
 VERSION AX608771.1 GI:28404323
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 REFERENCE 1
 AUTHORS Oi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.
 TITLE Novel serine protease genes related to DPPIV
 JOURNAL Patent: WO 0231134-A 48 18-APR-2002;
 Ferring BV (NL)
 FEATURES Location/Qualifiers
 source 1..22
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 0.7%; Score 22; DB 2; Length 22;
 Best Local Similarity 100.0%; Pred. No. 4.7; Mismatches 0; Indels 0; Gaps 0;
 Matches 22; Conservative 0;

QY 1880 CACATTCCTGTCATCAGTCA 1901
 |||
 DB 1 CACATTCCTGTCATCAGTCA 22

RESULT 103

AX608772/c
 LOCUS AX608772 22 bp DNA linear PAT 17-FEB-2003
 DEFINITION Sequence 49 from Patent WO0231134.
 ACCESSION AX608772
 VERSION AX608772.1 GI:28404324
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 REFERENCE 1
 AUTHORS Oi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.
 TITLE Novel serine protease genes related to DPPIV
 JOURNAL Patent: WO 0231134-A 49 18-APR-2002;
 Ferring BV (NL)
 FEATURES Location/Qualifiers
 source 1..22
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 0.7%; Score 22; DB 2; Length 22;
 Best Local Similarity 100.0%; Pred. No. 4.7; Mismatches 0; Indels 0; Gaps 0;
 Matches 22; Conservative 0;

QY 1972 TCAAGTCCTGAAGATGACCCAA 1993
 |||
 DB 22 TCAAGTCCTGAAGATGACCCAA 1

RESULT 104
 AR631301/c
 LOCUS AR631301 39 bp DNA linear PAT 14-FEB-2005
 DEFINITION Sequence 46 from patent US 6844180.
 ACCESSION AR631301
 VERSION AR631301.1 GI:59770960
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 39)
 AUTHORS Oi, S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.
 TITLE Serine protease genes related to DPPIV
 JOURNAL Patent: US 6844180-A 46 18-JAN-2005;
 Ferring BV;;
 FEATURES Location/Qualifiers
 source 1..39
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
 Query Match 0.7%; Score 22; DB 2; Length 39;
 Best Local Similarity 100.0%; Pred. No. 4.8; Mismatches 0; Indels 0; Gaps 0;
 Matches 22; Conservative 0;

QY 2839 ATTGCTGCTCTAAAGATGATAT 2860
 |||
 DB 39 ATTGCTGCTCTAAAGATGATAT 18

RESULT 105
 AX608769/c
 LOCUS AX608769 39 bp DNA linear PAT 17-FEB-2003
 DEFINITION Sequence 46 from Patent WO0231134.
 ACCESSION AX608769
 VERSION AX608769.1 GI:28404321
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;


```

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Qi,S., Akinsanya,K.O., Riviere,P.J. and Junien,J.L.
TITLE Novel serine protease genes related to dppiv
JOURNAL Patent: WO 0231134-A 46 18-APR-2002,
Ferring BV (NL)
FEATURES
source Location/Qualifiers
1..39
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 0.7%; Score 22; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2839 ATTCTGCTCTTAAAGTGATAT 2860
Db 39 ATTCTGCTCTTAAAGTGATAT 18
RESULT 106
CT030588 2250 bp mRNA linear VRT 12-OCT-2005
LOCUS Xenopus tropicalis finished cDNA, clone TTPA012113.
DEFINITION CT030588
VERSION CT030588.1 GI:77623488
KEYWORDS Xenopus tropicalis (Silurana tropicalis)
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 2250)
REFERENCE
AUTHORS Anaya,E., Ashurst,J.L., Bonfield,J.K., Croning,M.D.R., Davies,R.M.,
Francis,M.D., Garrett,N., Gilchrist,M.J., Graham,D.V.,
McLaren,S.R., Papalopulu,N., Rogers,J., Smith,J.C., Taylor,R.G.,
Voigt,J. and Zorn,A.M.
TITLE Direct Submission
JOURNAL Submitted (12-OCT-2005) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: trop@sanger.ac.uk
COMMENT Sanger Xenopus tropicalis EST/cDNA project.
This sequence is from a Xenopus Gene Collection (XGC) library, from
a library constructed by Nigel Garrett. cDNA was prepared from RNA
extracted from tadpoles. EcoRI-NotI cut cDNA was then ligated into
pCS107 with EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli DH10B.
FEATURES
source Location/Qualifiers
1..2250
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TTPA012113"
/clone_lib="XGC-tadpole"
/dev_stage="tadpole (stage 35-40)"
ORIGIN
Query Match 0.7%; Score 22; DB 11; Length 2250;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1073 ATGAGAAATGATGATCTGA 1094
Db 167 ATGAGAAATGATGATCTGA 188
RESULT 107
BD103383 3322 bp DNA linear PAT 27-AUG-2002
LOCUS Novel TTV-associated virus s-TTV.
DEFINITION BD103383
VERSION BD103383.1
KEYWORDS Novel TTV-associated virus s-TTV
SOURCE Patent: WO 0185771-A 10 15-NOV-2001;
EISAI CO LTD, KENJI ABE
REFERENCE
AUTHORS Abe,K.
TITLE Novel TTV-associated virus s-TTV
JOURNAL Patent: WO 0185771-A 10 15-NOV-2001;
EISAI CO LTD, KENJI ABE
COMMENT OS TTV CH71
PN WO 0185771-A/10
PD 15-NOV-2001
PF 11-MAY-2001 WO 2001JP003954
PR 11-MAY-2000 JP OOP 137894, 08-NOV-2000 JP OOP 340614 PI
PC C07K14/01, C12N15/51, G01N33/576, G01N33/569, C12Q1/68//C12P21/02
CC Novel TTV-associated virus s-TTV
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..3322
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
FEATURES
source
1..3322
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 0.7%; Score 22; DB 2; Length 3322;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 28 CGCTGCTACTGCCCGCGCTGCT 49
Db 115 CGCTGCTACTGCCCGCGCTGCT 136
RESULT 108
BD103384 3322 bp DNA linear PAT 27-AUG-2002
LOCUS Novel TTV-associated virus s-TTV.
DEFINITION BD103384
VERSION BD103384.1
KEYWORDS Novel TTV-associated virus s-TTV
SOURCE Patent: WO 0185771-A/10.
REFERENCE
AUTHORS Abe,K.
TITLE Novel TTV-associated virus s-TTV
JOURNAL Patent: WO 0185771-A 10 15-NOV-2001;
EISAI CO LTD, KENJI ABE
COMMENT OS TTV CH71
PN WO 0185771-A/10
PD 15-NOV-2001
PF 11-MAY-2001 WO 2001JP003954
PR 11-MAY-2000 JP OOP 137894, 08-NOV-2000 JP OOP 340614 PI
PC C07K14/01, C12N15/51, G01N33/576, G01N33/569, C12Q1/68//C12P21/02
CC Novel TTV-associated virus s-TTV
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..3322
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
FEATURES
source
1..3322
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 0.7%; Score 22; DB 2; Length 3322;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

BD103383
BD103383.1 GI:22648957
WO 0185771-A/9.
unidentified
ORGANISM unidentified
unclassified sequences.
1 (bases 1 to 3322)
REFERENCE
AUTHORS Abe,K.
TITLE Novel TTV-associated virus s-TTV
JOURNAL Patent: WO 0185771-A 9 15-NOV-2001;
EISAI CO LTD, KENJI ABE
COMMENT OS TTV CH71
PN WO 0185771-A/9
PD 15-NOV-2001
PF 11-MAY-2001 WO 2001JP003954
PR 11-MAY-2000 JP OOP 137894, 08-NOV-2000 JP OOP 340614 PI
PC C07K14/01, C12N15/51, G01N33/576, G01N33/569, C12Q1/68//C12P21/02
CC Novel TTV-associated virus s-TTV
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..3322
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
FEATURES
source
1..3322
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 0.7%; Score 22; DB 2; Length 3322;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 28 CGCTGCTACTGCCCGCGCTGCT 49
Db 115 CGCTGCTACTGCCCGCGCTGCT 136
RESULT 108
BD103384 3322 bp DNA linear PAT 27-AUG-2002
LOCUS Novel TTV-associated virus s-TTV.
DEFINITION BD103384
VERSION BD103384.1
KEYWORDS Novel TTV-associated virus s-TTV
SOURCE Patent: WO 0185771-A/10.
REFERENCE
AUTHORS Abe,K.
TITLE Novel TTV-associated virus s-TTV
JOURNAL Patent: WO 0185771-A 10 15-NOV-2001;
EISAI CO LTD, KENJI ABE
COMMENT OS TTV CH71
PN WO 0185771-A/10
PD 15-NOV-2001
PF 11-MAY-2001 WO 2001JP003954
PR 11-MAY-2000 JP OOP 137894, 08-NOV-2000 JP OOP 340614 PI
PC C07K14/01, C12N15/51, G01N33/576, G01N33/569, C12Q1/68//C12P21/02
CC Novel TTV-associated virus s-TTV
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..3322
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
FEATURES
source
1..3322
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 0.7%; Score 22; DB 2; Length 3322;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	28	CGCTGCTACTGCGCGCTGCT	49
Db	115	CGCTGCTACTGCGCGCTGCT	136
RESULT 109			
LOCUS	AB049608	3322 bp	linear
DEFINITION	TT virus gene for ORF1, ORF2, complete cds, isolate:s-TTV CH71.		
ACCESSION	AB049608		
VERSION	AB049608.1	GI:12060430	
KEYWORDS			
SOURCE	Torque teno virus (TTV)		
ORGANISM	Viruses; ssDNA viruses; Anellovirus.		
REFERENCE	1 (sites)		
AUTHORS	Abe, K. and Inami, T.		
TITLE	simian TT virus		
JOURNAL	Unpublished		
REFERENCE	2 (sites)		
AUTHORS	Abe, K. and Inami, T.		
TITLE	simian TT virus		
JOURNAL	Published Only in DataBase (2001)		
REFERENCE	3 (bases 1 to 3322)		
AUTHORS	Abe, K. and Inami, T.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-OCT-2000) Kenji Abe, National Institute of Infectious Diseases, Department of Pathology; Toyama 1-23-1, Shinjuku-ku, Tokyo 162-8640, Japan (E-mail: kenjiabe@nih.go.jp, Tel:81-3-5285-1111 (ex.2624), Fax:81-3-5285-1189)		
FEATURES	Location/Qualifiers		
source	1..3322		
	/organism="Torque teno virus"		
	/mol_type="genomic DNA"		
	/isolate="s-TTV CH71"		
	/specific_host="Pan troglodytes"		
	/db_xref="taxon:68887"		
	153..518		
	/codon_start=1		
	/product="ORF2"		
	/protein_id="BAB20603.1"		
	/db_xref="GI:12060431"		
	/translation="MSLWRPFAHDVTGLERNWYSCFRSHAMCGGCFIDHLNDLAS		
	RFGPPNRPFGAPQPPPIRFPALPAPEGERAPWRSGGGDAEGRGGDAPRRGGD		
	DGGDLADALLALEDAAE"		
	389..2572		
	/codon_start=1		
	/product="ORF1"		
	/protein_id="BAB20604.1"		
	/db_xref="GI:12060432"		
	/translation="MAFWRRRRRRRRAAARWRRRRRRRRRAAARAGRRRRVR		
	RRGRWRARRRRRRRRRRRRKLLIQQQPPYTRRCYIVGLYPLIMCGECFSKN		
	FATHSDDDICSVAGGGTATQFTLRILYDFORHLNWTASNNDDLDLARYLGATITF		
	FRHPDVFIVQIHITSPPPRDTEMGSIHFGMLMKHRLIIPSLKTRPGRKHRTVR		
	IGAPTFEDKWPQDCLDVILVYATAADFTFPFGSPDLTNHCFNQLGSDYNDI		
	LSLIPKLSKESKEIKYKNIAYNTKQTAHLKITDNCNTQTDITNKGNDQYK		
	GNTYDEINVRKDKGDFIKATISLALSNWKPPTNDLSEYTHGMYSILFSGRSNP		
	ELKGYTDVNCNPLVDKIGINVLWDCSKEDSKFDETSKVPTRDGLRMPGDSITYP		
	YASKETGDTAINTNARLILCEYTDPMLEKPDQPKFGVPVPSLNFGLRMPGDSITYP		
	IRQAKRYACLVHQHVEALITQSGFPAYHCDYKSAVLGVKYKFKWGSPIPHOVV		
	RNPRCTHSTQGGKRPKRSVQVTKKYNTPFTWHSNDTRGLFGDGLQRLMLQQLGL		
	NELYPTASKRPRKRTLAEGHQEPEPDSVDFKRLQLQQQPMLDSSQEEKKARESKA		
	QEEEAATLFEQLEQQRLGLFQLRCLAYQLVQVQKHAIPQALVQ"		
ORIGIN			
Query Match	0.7%;	Score 22;	DB 10; Length 3322;
Best Local Similarity	100.0%;	Pred. No. 5.3;	
Matches	22;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	28	CGCTGCTACTGCGCGCTGCT	49
Db	115	CGCTGCTACTGCGCGCTGCT	136

```

AP000861/c
LOCUS
DEFINITION Homo sapiens chromosome 11 clone CMB9-49N2 map liq25, WORKING DRAFT
SEQUENCE, 14 unordered pieces.
ACCESSION AP000861
VERSION AP000861.3 GI:8119009
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Homidae; Homo; Homo104012
REFERENCE
    1 (bases 1 to 104012)
    Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
    Fujiyama,A., Iada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
    Homo sapiens 104,012 genomic DNA of liq25
    Published Only in Database (1999)
REFERENCE
    2 (bases 1 to 104012)
    Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
    Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
    Direct Submission
    Submitted (13-DEC-1999) Masahira Hattori, The Institute of Physical
    and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
    Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
    Japan (E-mail:hattori@gsc.riken.go.jp,
    URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
    Fax:81-42-778-9924)
COMMENT
    On May 31, 2000 this sequence version replaced gi:7242895.
    ----- Genome Center
    Center: RIKEN Genomic Sciences Center (GSC)
    Center code: RIKEN
    Web site: http://hgp.gsc.riken.go.jp/
    Contact: hattori@gsc.riken.go.jp
    ----- Project Information
    Center project name: HumDraft11
    Center clone name: CMB9-49N2
    ----- Summary Statistics
    Sequencing vector: PCR products; 100% of reads
    Chemistry: Dye-terminator ET-amersham; 100% of reads
    Assembly program: Phrap; version 0.990329
    Consensus quality: 91543 bases at least Q40
    Consensus quality: 97504 bases at least Q50
    Consensus quality: 100932 bases at least Q20
    Insert size: 102712; sum-of-contigs
    Quality coverage: 4.48x in Q20 bases; sum-of-contigs
    -----
    NOTE: This is a 'working draft' sequence. It currently consists of
    14 contigs. The true order of the pieces is not known and their
    order in this sequence record is arbitrary. Gaps between the
    contigs are represented as runs N, but the exact sizes of the gaps
    are unknown. This record will be updated with the finished sequence
    as soon as it is available and the accession number will be
    preserved.
    1 20152 contig of 20152 bp in length
    20253 33930 contig of 13678 bp in length
    42313 58306 contig of 12182 bp in length
    58213 58225 contig of 11994 bp in length
    6326 74172 contig of 5847 bp in length
    74273 80325 contig of 6053 bp in length
    80426 85720 contig of 5295 bp in length
    85821 89487 contig of 3667 bp in length
    89588 93995 contig of 4408 bp in length
    94096 97810 contig of 3715 bp in length
    97911 100367 contig of 4579 bp in length
    100468 102265 contig of 1798 bp in length
    102366 104012 contig of 1647 bp in length
    Sequence updated (13-Mar-2000)
    Sequence updated (26-May-2000).
    * NOTE: This is a 'working draft' sequence. It currently
    * consists of 14 contigs. The true order of the pieces
    * is not known and their order in this sequence record is
    * arbitrary. Gaps between the contigs are represented as

```

```

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 20152: contig of 20152 bp in length
* 20253 33930: gap of 100 bp
* 33930 42313: contig of 13678 bp in length
* 42313 58213: gap of 100 bp
* 58213 6326: contig of 12182 bp in length
* 6326 74172: gap of 100 bp
* 74172 80325: contig of 11994 bp in length
* 80325 85720: gap of 100 bp
* 85720 89487: contig of 3667 bp in length
* 89487 93995: gap of 100 bp
* 93995 94096: contig of 4408 bp in length
* 94096 97810: gap of 100 bp
* 97810 97911: contig of 3715 bp in length
* 97911 100367: contig of 4579 bp in length
* 100367 100468: gap of 100 bp
* 100468 102265: contig of 1798 bp in length
* 102265 102366: gap of 100 bp
* 102366 104012: contig of 1647 bp in length.
FEATURES
    source
        1..104012
            Location/Qualifiers
                organism="Homo sapiens"
                mol_type="genomic DNA"
                db_xref="taxon:9606"
                chromosome="11"
                map="11Q25"
                clone="CMB9-49N2"
            1..20152
                /note="assembly_fragment"
            20253..33930
                /note="assembly_fragment"
            34031..46212
                /note="assembly_fragment"
            46313..58306
                /note="assembly_fragment"
            58407..68225
                /note="assembly_fragment"
            68326..74172
                /note="assembly_fragment clone_end:T7 vector_side:right"
            74273..80325
                /note="assembly_fragment"
            80426..85720
                /note="assembly_fragment"
            85821..89487
                /note="assembly_fragment"
            89588..93995
                /note="assembly_fragment"
            94096..97810
                /note="assembly_fragment"
            97911..100367
                /note="assembly_fragment"
            100468..102265
                /note="assembly_fragment"
            102366..104012
                /note="assembly_fragment"
ORIGIN
Query Match 0.7%; Score 22; DB 12; Length 104012;
Best Local Similarity 100.0%; Pred. NO. 5.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

* consists of 11 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 17900: contig of 17900 bp in length
 17901 18000: gap of 100 bp
 18001 34310: contig of 16310 bp in length
 34311 34411: gap of 100 bp
 34411 48225: contig of 13815 bp in length
 48226 48325: gap of 100 bp
 48326 62336: contig of 14011 bp in length
 62337 62437: gap of 100 bp
 62437 76488: contig of 14052 bp in length
 76489 76588: gap of 100 bp
 76589 84448: contig of 7860 bp in length
 84449 84548: gap of 100 bp
 84549 92441: contig of 7893 bp in length
 92442 92541: gap of 100 bp
 92542 99422: contig of 6881 bp in length
 99423 103359: contig of 100 bp
 103360 103459: gap of 100 bp
 103460 104651: contig of 1192 bp in length
 104652 104751: gap of 100 bp
 104752 105350: contig of 599 bp in length.

FEATURES

Location/Qualifiers
 1. .105350
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q25"
 /clone="CMB9-2G15"
 1. .17900
 /note="assembly_fragment"
 18001. .34310
 /note="assembly_fragment"
 34411. .48225
 /note="assembly_fragment"
 48326. .62336
 /note="assembly_fragment"
 62437. .76488
 /note="assembly_fragment"
 76589. .84448
 /note="assembly_fragment"
 84549. .92441
 /note="assembly_fragment"
 92542. .99422
 /note="assembly_fragment clone_end:77 vector_side:right"
 99523. .103359
 /note="assembly_fragment"
 103460. .104651
 /note="assembly_fragment"
 104752. .105350
 /note="assembly_fragment clone_end:SP6 vector_side:left"

ORIGIN

Query Match 0.7%; Score 22; DB 12; Length 105350;
 Best Local Similarity 100.0%; Pred.No. 5.7;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3098 ATAAGAAATTACTAAAAAAA 3119
 Db 88539 ATAAGAAATTACTAAAAAAA 88518

RESULT 113
 AP008208_040/c
 WPCOMMENT

Sequence split into 360 fragments LOCUS AP008208 Accession AP008208

QY 3098 ATAAGAAATTACTAAAAAAA 3119
 Db 84413 ATAAGAAATTACTAAAAAAA 84392

RESULT 112

AP000661/c 105350 bp DNA linear HTG 30-MAY-2000
 LOCUS Homo sapiens chromosome 11 clone CMB9-2G15 map 11q25, WORKING DRAFT
 DEFINITION SEQUENCE, 11 unordered pieces.

ACCESSION AP000661

VERSION HTG; HTGS PHASE1; HTGS_DRAFT.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE

AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
 Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

TITLE

Published Only in Database (1999)

JOURNAL

1 (bases 1 to 105350)

REFERENCE

AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
 Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 Direct Submission
 Submitted (04-NOV-1999) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
 Japan [E-mail:hattori@gsc.riken.go.jp,
 URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
 Fax:81-42-778-9924]
 On May 31, 2000 this sequence version replaced gi:6997535.

COMMENT

----- Genome Center
 Center: RIKEN Genomic Sciences Center (GSC)
 Center code: RIKEN
 Web site: http://hgp.gsc.riken.go.jp/
 Contact: hattori@gsc.riken.go.jp
 ----- Project Information
 Center project name: HumDraft11
 Center clone name: CMB9-2G15
 ----- Summary Statistics
 Sequencing vector: PCR products; 100% of reads
 Chemistry: Dye-terminator ET-amersham; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 101290 bases at least Q40
 Consensus quality: 103227 bases at least Q30
 Consensus quality: 103986 bases at least Q20
 Insert size: 104350; sum-of-contigs
 Quality coverage: 8.12x in Q20 bases; sum-of-contigs

 NOTE: This is a 'working draft' sequence. It currently consists of
 11 contigs. The true order of the pieces is not known and their
 order in this sequence record is arbitrary. Gaps between the
 contigs are represented as runs N, but the exact sizes of the gaps
 are unknown. This record will be updated with the finished sequence
 as soon as it is available and the accession number will be
 preserved
 1 17900 contig of 17900 bp in length
 18001 34310 contig of 16310 bp in length
 34411 48225 contig of 13815 bp in length
 48326 62336 contig of 14011 bp in length
 62437 76488 contig of 14052 bp in length
 76589 84448 contig of 7860 bp in length
 84549 92441 contig of 7893 bp in length
 92542 99422 contig of 6881 bp in length
 99523 103359 contig of 100 bp in length
 103460 104651 contig of 1192 bp in length
 104752 105350 contig of 599 bp in length

Sequence updated (06-Jan-2000)
 Sequence updated (26-May-2000).

* NOTE: This is a 'working draft' sequence. It currently

Fragment Name	Begin	End
AP008208_000	110000	7310000
AP008208_001	100001	7300001
AP008208_002	200001	7410000
AP008208_003	300001	7510000
AP008208_004	400001	7600001
AP008208_005	500001	7700001
AP008208_006	600001	7800001
AP008208_007	700001	7900001
AP008208_008	800001	8010000
AP008208_009	900001	8110000
AP008208_010	1000001	8210000
AP008208_011	1100001	8310000
AP008208_012	1200001	8400001
AP008208_013	1300001	8510000
AP008208_014	1400001	8610000
AP008208_015	1500001	8710000
AP008208_016	1600001	8810000
AP008208_017	1700001	8900001
AP008208_018	1800001	9000001
AP008208_019	1900001	9100000
AP008208_020	2000001	9200001
AP008208_021	2100001	9300001
AP008208_022	2200001	9400001
AP008208_023	2300001	9500001
AP008208_024	2400001	9600001
AP008208_025	2500001	9700001
AP008208_026	2600001	9800001
AP008208_027	2700001	9900001
AP008208_028	2800001	10010000
AP008208_029	2900001	10110000
AP008208_030	3000001	10210000
AP008208_031	3100001	10310000
AP008208_032	3200001	10400001
AP008208_033	3300001	10500001
AP008208_034	3400001	10600001
AP008208_035	3500001	10700001
AP008208_036	3600001	10800001
AP008208_037	3700001	10900001
AP008208_038	3800001	11000001
AP008208_039	3900001	11100001
AP008208_040	4000001	11200001
AP008208_041	4100001	11300001
AP008208_042	4200001	11400001
AP008208_043	4300001	11500001
AP008208_044	4400001	11600001
AP008208_045	4500001	11700001
AP008208_046	4600001	11800001
AP008208_047	4700001	11900001
AP008208_048	4800001	12000001
AP008208_049	4900001	12100001
AP008208_050	5000001	12200001
AP008208_051	5100001	12300001
AP008208_052	5200001	12400001
AP008208_053	5300001	12500001
AP008208_054	5400001	12600001
AP008208_055	5500001	12700001
AP008208_056	5600001	12800001
AP008208_057	5700001	12900001
AP008208_058	5800001	13000001
AP008208_059	5900001	13100001
AP008208_060	6000001	13200001
AP008208_061	6100001	13300001
AP008208_062	6200001	13400001
AP008208_063	6300001	13500001
AP008208_064	6400001	13600001
AP008208_065	6500001	13700001
AP008208_066	6600001	13800001
AP008208_067	6700001	13900001
AP008208_068	6800001	14000001
AP008208_069	6900001	14100001
AP008208_070	7000001	14200001
AP008208_071	7100001	14300001
AP008208_072	7200001	14400001
AP008208_073	7300001	14500001
AP008208_074	7400001	14600001
AP008208_075	7500001	14700001
AP008208_076	7600001	14800001
AP008208_077	7700001	14900001
AP008208_078	7800001	15000001
AP008208_079	7900001	15100001
AP008208_080	8000001	15200001
AP008208_081	8100001	15300001
AP008208_082	8200001	15400001
AP008208_083	8300001	15500001
AP008208_084	8400001	15600001
AP008208_085	8500001	15700001
AP008208_086	8600001	15800001
AP008208_087	8700001	15900001
AP008208_088	8800001	16000001
AP008208_089	8900001	16100001
AP008208_090	9000001	16200001
AP008208_091	9100001	16300001
AP008208_092	9200001	16400001
AP008208_093	9300001	16500001
AP008208_094	9400001	16600001
AP008208_095	9500001	16700001
AP008208_096	9600001	16800001
AP008208_097	9700001	16900001
AP008208_098	9800001	17000001
AP008208_099	9900001	17100001
AP008208_100	10000001	17200001
AP008208_101	10100001	17300001
AP008208_102	10200001	17400001
AP008208_103	10300001	17500001
AP008208_104	10400001	17600001
AP008208_105	10500001	17700001
AP008208_106	10600001	17800001
AP008208_107	10700001	17900001
AP008208_108	10800001	18000001
AP008208_109	10900001	18100001
AP008208_110	11000001	18200001
AP008208_111	11100001	18300001
AP008208_112	11200001	18400001
AP008208_113	11300001	18500001
AP008208_114	11400001	18600001
AP008208_115	11500001	18700001
AP008208_116	11600001	18800001
AP008208_117	11700001	18900001
AP008208_118	11800001	19000001
AP008208_119	11900001	19100001
AP008208_120	12000001	19200001
AP008208_121	12100001	19300001
AP008208_122	12200001	19400001
AP008208_123	12300001	19500001
AP008208_124	12400001	19600001
AP008208_125	12500001	19700001
AP008208_126	12600001	19800001
AP008208_127	12700001	19900001
AP008208_128	12800001	20000001
AP008208_129	12900001	20100001
AP008208_130	13000001	20200001
AP008208_131	13100001	20300001
AP008208_132	13200001	20400001
AP008208_133	13300001	20500001
AP008208_134	13400001	20600001
AP008208_135	13500001	20700001
AP008208_136	13600001	20800001
AP008208_137	13700001	20900001
AP008208_138	13800001	21000001
AP008208_139	13900001	21100001
AP008208_140	14000001	21200001
AP008208_141	14100001	21300001
AP008208_142	14200001	21400001
AP008208_143	14300001	21500001
AP008208_144	14400001	21600001

```

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens genomic DNA
TITLE
JOURNAL
Published Only in Database (2000)
REFERENCE
AUTHORS
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
TITLE
JOURNAL
Submitted (07-APR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Sushiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsr.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Nov 5, 2001 this sequence version replaced gi:11138088.
COMMENT
FEATURES
Location/Qualifiers
1..117345
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-700F16"
source

ORIGIN
Query Match 0.7%; Score 22; DB 5; Length 117345;
Best Local Similarity 100.0%; Pred. No. 5.7; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

Qy 3098 ATAGGAATTACTTAAAAA 3119
Db 4058 ATAGGAATTACTTAAAAA 4079

RESULT 115
AL160263 117445 bp DNA linear HTG 10-JUL-2001
LOCUS Homo sapiens chromosome 6 clone RP1-206M12, 22 unordered pieces.
DEFINITION
ACCESSION AL160263
VERSION AL160263.6 GI:9863592
KEYWORDS HTG; HTGS PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS
Sims,S.
TITLE
JOURNAL
Direct Submission
Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Aug 21, 2000 this sequence version replaced gi:8977651.
COMMENT
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: du206M12
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: M13; M7815; 4% of reads
Sequencing vector: plasmid; L08752; 95% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 104927 bases at least Q40
Consensus quality: 109617 bases at least Q30
Consensus quality: 112114 bases at least Q20
Insert size: 115345; sum-of-contigs
Insert size: 179782; 18.8% error; agarose-fp
Quality coverage: 4.71x in Q20 bases; sum-of-contigs Quality

14500001 14510000
14600001 14710000
14700001 14810000
14800001 14910000
14900001 15010000
15000001 15110000
15100001 15210000
15200001 15310000
15300001 15410000
15400001 15510000
15500001 15610000
15600001 15710000
15700001 15810000
15800001 15910000
15900001 16010000
16000001 16110000
16100001 16210000
16200001 16310000
16300001 16410000
16400001 16510000
16500001 16610000
16600001 16710000
16700001 16810000
16800001 16910000
16900001 17010000
17000001 17110000
17100001 17210000
17200001 17310000
17300001 17410000
17400001 17510000
17500001 17610000
17600001 17710000
17700001 17810000
17800001 17910000
17900001 18010000
18000001 18110000
18100001 18210000
18200001 18310000
18300001 18410000
18400001 18510000
18500001 18610000
18600001 18710000
18700001 18810000
18800001 18910000
18900001 19010000
19000001 19110000
19100001 19210000
19200001 19310000
19300001 19410000
19400001 19510000
19500001 19610000
19600001 19710000

Query Match 0.7%; Score 22; DB 4; Length 110000;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 986 AAGAAGATTGATGATATTC 1007
Db 46307 AAGAAGATTGATGATATTC 46286

RESULT 114
AP001775
LOCUS Homo sapiens genomic DNA, chromosome 11q clone:RP11-700F16,
DEFINITION complete sequence.
ACCESSION AP001775
VERSION AP001775.4 GI:16751482
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

us-10-825-632-2.olig.rge

coverage: 3.68x in Q20 bases, agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 22 contigs. The order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence, as soon as it is available and the accession number will be preserved.

1 4694: contig of 4694 bp in length
4695 4794: gap of 100 bp
4795 7828: contig of 3034 bp in length
7829 7929: gap of 100 bp
7929 10684: contig of 2756 bp in length
10685 10784: gap of 100 bp
10785 13948: contig of 3164 bp in length
13949 14048: gap of 100 bp
14049 17288: contig of 3240 bp in length
17289 17388: gap of 100 bp
17389 21695: contig of 4307 bp in length
21696 21795: gap of 100 bp
21796 24617: gap of 2822 bp in length
24618 29336: contig of 4619 bp in length
29337 33478: gap of 100 bp
33479 33578: contig of 4042 bp in length
33579 36895: contig of 3317 bp in length
36896 36995: gap of 100 bp
36996 39662: contig of 2566 bp in length
39663 44842: gap of 100 bp
44843 47261: contig of 5080 bp in length
47262 47362: contig of 2420 bp in length
47363 50478: gap of 100 bp
50479 50578: contig of 3117 bp in length
50579 93459: gap of 100 bp
93460 93559: gap of 100 bp
93560 96033: contig of 2474 bp in length
96034 96133: gap of 100 bp
96134 100979: contig of 4846 bp in length
100980 101079: gap of 100 bp
101080 103587: contig of 2508 bp in length
103588 106442: contig of 2755 bp in length
106443 106542: gap of 100 bp
106543 109973: contig of 3431 bp in length
109974 110074: gap of 100 bp
110075 114633: contig of 4561 bp in length
114634 114735: gap of 100 bp
114735 117445: contig of 2711 bp in length.

FEATURES

source

1..117445
Location/Qualifiers
organism="Homo sapiens"
mol_type="genomic DNA"
db_xref="taxon:9606"
chromosome="6"
clone="RP1-206M12"
clone_lib="RPC1-1"
1..4694
note="assembly_fragment:00720
fragment_chain:1"
4795..7828
note="assembly_fragment:01211
fragment_chain:1"
7929..10684
note="assembly_fragment:00482
fragment_chain:1"
10785..13948
note="assembly_fragment:01140
fragment_chain:1"

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature 14049..17288
/note="assembly_fragment:00231
fragment_chain:2"
misc_feature 17389..21695
/note="assembly_fragment:00129
fragment_chain:2"
misc_feature 21796..24617
/note="assembly_fragment:01301
fragment_chain:3"
misc_feature 24718..29336
/note="assembly_fragment:00659
fragment_chain:3"
misc_feature 29437..33478
/note="assembly_fragment:01946
fragment_chain:4"
misc_feature 33579..36895
/note="assembly_fragment:00111
fragment_chain:4"
misc_feature 36996..39561
/note="assembly_fragment:00205"
39662..44741
/note="assembly_fragment:00242"
44842..47261
misc_feature 47362..50478
/note="assembly_fragment:00673"
50579..93459
/note="assembly_fragment:00870"
93560..96033
/note="assembly_fragment:01072"
96134..100979
misc_feature 101079..103587
/note="assembly_fragment:01517"
103688..106442
misc_feature 106543..109973
/note="assembly_fragment:01608"
110074..114633
misc_feature 110074..114633
/note="assembly_fragment:01719"
114735..117445
misc_feature 114735..117445
/note="assembly_fragment:01865"
/note="assembly_fragment:01866"

ORIGIN

Query Match 0.7%; Score 22; DB 12; Length 117445;
Best Local Similarity 100.0%; Pred. No. 5.7; 0; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0;

Qy 3099 TAAGATTACTAAAAA 3120
|||||

Db 52142 TAAGATTACTAAAAA 52163
|||||

RESULT 116

CR749173
LOCUS CR749173 143028 bp DNA linear HTG 03-MAR-2005
DEFINITION Danio rerio clone.DKEYP-73C1, 5 unordered pieces.
ACCESSION CR749173
VERSION CR749173.10 GI:60495570
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 Phases 1 to 143028
Phases 1 to 143028
DIRECT SUBMISSION
Submitted (02-MAR-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Mar 3 2005 this sequence version replaced gi:56783567.
----- Genome Center

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfsh-help@sanger.ac.uk
----- Project Information
Center project name: zkp73C1
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 140742 bases at least Q40
Consensus quality: 141127 bases at least Q30
Consensus quality: 141713 bases at least Q20
Insert size: 142628; sum-of-contigs
Insert size: 196437; 2.5% error; agarose-fp
Quality coverage: 7.92x in Q20 bases; sum-of-contigs Quality
Coverage: 6.95x in Q20 bases; agarose-fp

HTG
Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 167262)
Clark, S
Direct Submission
Submitted (29-JAN-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clonerequests@sanger.ac.uk
zfsh-help@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Jan 29, 2005 this sequence version replaced gi:56310288.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfsh-help@sanger.ac.uk

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 123412: contig of 123412 bp in length
* 123413 123512: gap of 100 bp
* 123513 135554: contig of 12042 bp in length
* 135554 135655: gap of 100 bp
* 135655 138272: contig of 2618 bp in length
* 138272 138373: gap of 100 bp
* 138373 140636: contig of 2264 bp in length
* 140636 140737: gap of 100 bp
* 140737 143028: contig of 2292 bp in length.
----- Location/Qualifiers
1. .143028
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone_lib="DKEYP-73C1"
/clone_lib="DanioKeyPilot"
1. .123412
/note="assembly fragment:00549
fragment_chain:1"
123513_135554
/note="assembly fragment:01662
fragment_chain:1"
135655_138272
/note="assembly fragment:01725
fragment_chain:2"
138373_140636
/note="assembly fragment:01743
fragment_chain:2"
140737_143028
/note="assembly fragment:01700"

FEATURES

source
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature

ORIGIN

Query Match 0.7%; Score 22; DB 12; Length 143028;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3051 ATGCTCAAAATCAATGATACA 3072
|||||
DB 36416 ATGCTCAAAATCAATGATACA 36437
|||||

RESULT 117

CR847930 CR847930 167262 bp DNA linear VRT 29-JAN-2005
LOCUS Zebrafish DNA sequence from clone DKEY-148H19 in linkage group 19,
DEFINITION complete sequence.
ACCESSION CR847930
VERSION CR847930.5 GI:58331654

FEATURES

source
1. .167262
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-148H19"
/clone_lib="DanioKey"

ORIGIN

Query Match 0.7%; Score 22; DB 11; Length 167262;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3051 ATGCTCAAAATCAATGATACA 3072
|||||
DB 166541 ATGCTCAAAATCAATGATACA 166562
|||||

RESULT 118

AC158877/c AC158877 169540 bp DNA linear HTG 24-MAR-2005
LOCUS Bos taurus clone CH240-22115, WORKING DRAFT SEQUENCE, 23 unordered
DEFINITION pieces.
ACCESSION AC158877
VERSION AC158877

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Emi, EMBL; Swi, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep Clone-derived
zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.
Repeat names beginning 'Dr' were identified by the Recon repeat
discovery system (Zhihong Bao and Sean Eddy, submitted), and those
beginning 'drr' were identified by Rick Waterman (Stephen Johnson
lab, WashU). For further information see
http://www.sanger.ac.uk/projects/D_rerio/fishmask.shtml DKEY-148H19
is from a Zebrafish BAC library
VECTOR: pindigobAC-5.

```

VERSION
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE
ORGANISM
Bos taurus (cattle)

REFERENCE
AUTHORS
AC158877.1 GI:61740749
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 169540)
Murny,D., Adams,C., Agbai II,O., Allen,C., Alsbrooks,S., Archer,P.,
Arredondo,H., Bandaruaike,D., Bangura,L., Beltran,B., Beltran,R.,
Beraducci,A., Biswal,K., Blyth,P., Bonham,H., Buhay,C., Burch,P.,
Cadorese,A., Cardenas,V., Carter,K., Chen,A., Chen,G., Chen,R.,
Chacko,J., Chaurour,M., Chavez,D., Chen,A., Chen,G., Chen,R.,
Curry,K.-I., Chu,J., Clerc,K., Cockrell,R., Coyle,M., Cree,A.,
Anda,C., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H.,
Donlin,J., McCauley,S., Dugan-Rocha,S., Dunn,A., Durbin,K.,
Dziudza,D., Egan,A., Escotto,M., Espinosa,V., Eugene,C., Fa,M.,
Fernandez,S., Fernando,P., Flagg,N., Forbes,L., Foster,P.,
Fowler,G., Fu,Q., Fuh,E., Garcia,A., Garcia,R., Garner,T.,
Gaskin,C., Gench,S., Ghose,S., Gill,R., Gonzalez,D.,
Gonzalez-Garay,M., Guevara,W., Holder,M., Haaland,W., Haeblerlen,K.,
Hall,B., Hamid,H., Hamilton,K., Harbes,B., Harris,R., Havlak,P.,
Hines,S., Hitchens,M., Hodgson,A., Hogue,M., Hollins,B.,
Howell,L.T., Hulyk,S., Hume,J., Imo,K., Jackson,A., Jackson,L.,
Jacob,L., Jiang,H., Johnson,B., Johnson,R., Kalafus,K., Kelly,S.,
Keys,T., Khan,Z., King,L., Kovar,C., Kowis,A., Kowis,C., Lara,F.,
Leal,S., Lee,K., Lee,S., LeGall,F.I., Lemon,S., Lewis,L., Li,B.,
Li,Y., Li,Z., Linnell,M., Liu,W., Liu,Y.-S., Liu,Y., Liyanage,D.,
London,P., Lopez,J., Lorenzshew,L., Lozano,R., Luk,T., Madue,R.,
Maheshwari,M., Mahoney,C., Malloy,K., Mansouri,D., Martinez,E.,
McClalland,H., McPherson,J., Mercado,C., Metzker,M.,
Milosavljevic,A., Minja,E., Morgan,M., Morris,S.,
Murray,D., Nazareth,L., Ngo,D., Nguyen,N., Norwig-Eastaugh,E.,
Ottawa,K., Nwaokemele,O., Oregon,M., Ochi-Okeke,C., Oden,E.,
Okwuonu,G., Okwuonu,K., Parker,D., Pasternak,S., Patel,B.,
Patel,V., Paul,H., Perez,A., Perez,L., Petrosino,J., Pham,T.,
Primus,E., Pu,L.-L., Puazo,M., Qian,X., Quinn,A., Quiroz,J.,
Rabata,D., Rachlin,E., Reich,R., Reiz,R., Reuter,M., Richards,S.,
Rives,C., Rodriguez,F., Rojas,A., Ruiz,S.J., Sana,M., Sanders,W.,
Santibanez,J., Santos,R., Savary,G., Scherer,S., Shen,H., Shen,Y.,
Sisson,I., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R.,
Svatek,A., Taylor,E., Taylor,T., Thomas,N., Thorne,R., Thornton,R.,
Trejos,Z., Umami,K., Vargo,C., Verdusco,D., Villalana,D., Virk,D.,
Volkov,A., Waldron,J., Walker,B., Wang,Q., Wang,S., Warren,J.,
Wei,X., Wheeler,D., Williams,G., Williams,R., Worley,K., Wright,R.,
Wu,J., Yabub,S., Yan,K., Yaun,Y., Yu,P., Zhang,J., Zhang,L.,
Zhang,Z., Zhou,J., Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 169540)
Worley,K.C.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help.tmc.edu
----- Project Information
Center project name: FBAB
Center clone name: CH240-22115
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 171082 bases at least Q40
Consensus quality: 172961 bases at least Q30
Consensus quality: 175246 bases at least Q20
Estimated insert size: 180839; sum-of-contigs estimation

```

Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently consists of 23 contigs. The true order of the pieces is not known and their order in this sequence is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

Location/Qualifiers	Contig	Length
1	3643:	contig of 3643 bp in length
3644	3743:	gap of unknown length
3744	13031:	contig of 9288 bp in length
13032	13131:	gap of unknown length
13132	16258:	contig of 3127 bp in length
16259	16358:	gap of unknown length
16359	24536:	contig of 8078 bp in length
24537	24536:	gap of unknown length
24537	27320:	contig of 2784 bp in length
27321	27320:	gap of unknown length
27321	30906:	contig of 3486 bp in length
30907	30906:	gap of unknown length
31007	39018:	contig of 8012 bp in length
39019	39118:	gap of unknown length
39119	43186:	contig of 4068 bp in length
43187	43286:	gap of unknown length
43287	54202:	contig of 10916 bp in length
54203	54302:	gap of unknown length
54303	56596:	contig of 2294 bp in length
56597	56596:	gap of unknown length
56697	59280:	contig of 2584 bp in length
59281	59380:	gap of unknown length
59381	64409:	contig of 5029 bp in length
64410	64509:	gap of unknown length
64510	78550:	contig of 14041 bp in length
78551	78550:	gap of unknown length
78551	81198:	contig of 2548 bp in length
81199	81298:	gap of unknown length
81299	86477:	contig of 5179 bp in length
86478	86577:	gap of unknown length
86578	101411:	contig of 14834 bp in length
101412	101511:	gap of unknown length
101512	104847:	contig of 3336 bp in length
104848	104947:	gap of unknown length
104948	111323:	contig of 6376 bp in length
111324	111423:	gap of unknown length
111424	125250:	contig of 13827 bp in length
125251	125250:	gap of unknown length
125351	135463:	contig of 10113 bp in length
135464	135463:	gap of unknown length
135564	150580:	contig of 14987 bp in length
150581	150580:	gap of unknown length
150581	157250:	contig of 7000 bp in length
157251	157250:	gap of unknown length
157251	169540:	contig of 11790 bp in length.

[illegible]

GFFPQILGSRKRGAPHRSPLS"
 25102..25323
 /gene="OSJNBA0073A21.3"
 25102..25323
 /gene="OSJNBA0073A21.3"
 /note="transposable elements TNP2 protein-like"
 30082..31019
 /gene="OSJNBA0073A21.4"
 30082..31019
 /gene="OSJNBA0073A21.4"
 /note="probably inactive due to 5' exon missing in CDS
 pseudogene, putative cyst nematode resistance protein"
 pseudogene
 complement(join(32626..32682,32727..32940,33036..33142))
 /gene="OSJNBA0073A21.5"
 complement(join(32626..32682,32727..32940,33036..33142))
 /gene="OSJNBA0073A21.5"
 /note="hypothetical ORF
 predicted by GlimmerM
 this category is not included in IRGSP standard"
 complement(37925..38191)
 /gene="OSJNBA0073A21.6"
 complement(37925..38191)
 /gene="OSJNBA0073A21.6"
 /note="hypothetical ORF
 predicted by GlimmerM
 this category is not included in IRGSP standard"
 complement(38706..41739)
 /gene="OSJNBA0073A21.7"
 complement(join(<38706..39166,39255..40218,40327..41444,
 41538..>41739))
 /gene="OSJNBA0073A21.7"
 /note="supported by full-length cDNA(s) : AKI19397"
 complement(38706..41739)
 /gene="OSJNBA0073A21.7"
 /note="contains full-length cDNA(s) : AKI19397
 non-coding transcript
 probably inactive due to including stop codon(s) in CDS"
 43643..49287
 /gene="OSJNBA0073A21.8-1"
 43643..49287
 /gene="OSJNBA0073A21.8-2"
 /note="contains full-length cDNA(s) : AK066680
 non-coding transcript
 probably inactive due to including stop codon(s) in CDS"
 join(42848..42972,43128..43311,44727..44834,45075..45159,
 45455..45690,45916..46089,46176..46223,46325..46384,
 46481..46567,46673..46761,46875..46966,47776..47837,
 47937..48086,48173..48388)
 /gene="OSJNBA0073A21.8-1"
 join(<42848..42972,43128..43311,44727..44834,45075..45159,
 45455..45690,45916..46089,46176..46223,46325..46384,
 46481..46567,46673..46761,46875..46966,47776..47837,
 47937..48086,48173..48388)
 /gene="OSJNBA0073A21.8-1"
 /note="contains EST(s) : AU173808(R3811),AU173809(R3811)
 contains full-length cDNA(s) : AK062235"
 /codon_start=1
 /product="putative cytochrome P450"
 /protein_id="BAD26183.1"
 /db_xref="GI:49388966"
 /translation="MAITATAAAATPHQADASPRHAACPALGRBRLPVVRQ
 SSVVDKPKSKRGGLDNVSLNLSGGSLGMPVAGVAVDLPGRPLPSLVNWF
 EBGVYKLAFLPKAFVVSPIVARHILNENAFYKQVLAEILKPKGLIDDL
 TQKRRKVIPTGFHALPIDAMGVFTKCSERTFKLEELIERGEKVTYDLAEF
 SNLADILIGLVGFDFDSVTKESPVIKAVYGTLPFAEHRSTFYIPVNNLEIRWIP
 RQRKFHSLDKLVINDCLSLKNAKETRQEADEVKQORDYSSLKQASLIRFVDMRGA

DVDDQLRDLMTMLIAGHETTAVALTWVSFLLAQNFSKMRQAQAEVDSVLSNETINV
 DQKKLEYIRLIIVAEALRLYPOPLLRRLRPDKLPGGYNGAKEGVIEIPACTDIFLS
 IYNLHRSFYFWRDPDEPERFSVPKDESIKGMAGDPDRSPGMYPNEILADFAFL
 PFGGPRKCVGDQFALLESTVALALLQKFDVLRGSPDEVNTGATHTKSGLMCR
 VRRRT"
 43094..49322
 /gene="OSJNBA0073A21.8-3"
 join(<43094..43311,44727..44834,45075..45159,45455..45690,
 45916..46089,46176..46223,46301..46384,46481..46567,
 46673..46761,46875..46966,47776..47837,47937..48086,
 48173..48412,48903..48932)
 /gene="OSJNBA0073A21.8-3"
 /note="supported by full-length cDNA(s) : AKI00596"
 43094..49322
 /gene="OSJNBA0073A21.8-3"
 /note="contains full-length cDNA(s) : AKI00596
 non-coding transcript
 probably inactive due to including stop codon(s) in CDS"
 55652..56799
 /gene="OSJNBA0073A21.9"
 55652..56799
 /gene="OSJNBA0073A21.9"
 /note="supported by full-length cDNA(s) : AK068418"
 55672..56361
 /gene="OSJNBA0073A21.9"
 /note="contains full-length cDNA(s) : AK068418"
 /codon_start=1
 /product="unknown protein"
 Query Match 0.7%; Score 22; DB 4; Length 170051;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 986 AAGAAGAAATTGATAGATATTC 1007
 DB 92973 AAGAAGAAATTGATAGATATTC 92952
 AC136070 180005 bp DNA linear HTG 29-OCT-2002
 Rattus norvegicus clone CH230-321022, *** SEQUENCING IN PROGRESS
 *** 66 unordered pieces.
 AC136070.1 GI:24417941
 HTG: HTGS PHASE1
 Rattus norvegicus
 Rattus norvegicus (Norway rat)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Sciurognathi; Muroidea; Muridae; Murinae; Rattus;
 1 (bases 1 to 180005)
 Muzny, D. Marie, Metzker, M. Lee, Abranzon, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, R., Bunay, C., Burck, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, K., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Evans, C., Evans, C.A., Falle, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

	Karpthy,S., Kelly,S., Khan,Z., King,L., Kovar,C.,	7461	8749: contig of 1289 bp in length
	Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,	8750	8849: gap of unknown length
	Liu,J., Liu,Y., London,P., Longacre,S., Lopez,J.,	8850	9996: contig of 1147 bp in length
	Lorenuhewa,L., Louisseg,H., Lozado,R.J., Lu,X., Ma,J.,	9997	10096: gap of unknown length
	Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,	10097	11833: contig of 1737 bp in length
	Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,	11834	11833: gap of unknown length
	Mawhney,S., McLeod,M., McNeill,T., Meenen,E., Milosavljevic,A.,	11934	13397: contig of 1464 bp in length
	Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K.,	13398	13497: gap of unknown length
	Morris,S., Muidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D.,	13498	14947: contig of 1450 bp in length
	Newton,N., Nguyen,N., Norris,S., Nwaokemele,O., Okunnu,G.,	14948	15047: gap of unknown length
	Olarnpunagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H.,	15048	16173: contig of 1126 bp in length
	Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Poindexter,A.,	16174	16273: gap of unknown length
	Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E.,	16274	18356: contig of 2083 bp in length
	Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y.,	18357	18456: gap of unknown length
	Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A.,	18457	20492: contig of 2036 bp in length
	Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savory,G., Scherer,S.,	20493	20922: gap of unknown length
	Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A.,	20922	22782: contig of 2190 bp in length
	Song,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E.,	22783	22882: gap of unknown length
	Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A.,	22883	23904: contig of 1022 bp in length
	Stavak,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D.,	23905	24004: gap of unknown length
	Tingey,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S.,	24005	26411: contig of 2407 bp in length
	Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J.,	26412	26511: gap of unknown length
	Warren,R., Wei,X., White,F., Williams,G., Wilson,R., Wleczyk,R.,	28084	28183: gap of unknown length
	Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S.,	28184	29417: contig of 1234 bp in length
	Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X.,	29418	29517: gap of unknown length
	Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R.,	29518	31138: contig of 1621 bp in length
	Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.	31139	31238: gap of unknown length
	Direct Submission	31239	32835: contig of 1597 bp in length
	Unpublished	32836	32935: gap of unknown length
	2 (bases 1 to 180005)	32936	34816: contig of 1881 bp in length
	Rat Genome Sequencing Consortium.	34817	36426: contig of 1510 bp in length
	Direct Submission	36427	36526: gap of unknown length
	Submitted (29-OCT-2002) Human Genome Sequencing Center, Department	36527	38231: contig of 1705 bp in length
	of Molecular and Human Genetics, Baylor College of Medicine, One	38232	38331: gap of unknown length
	Baylor Plaza, Houston, TX 77030, USA	38332	39365: contig of 1034 bp in length
	----- Genome Center	39366	39465: gap of unknown length
	Center: Baylor College of Medicine	39466	41078: contig of 1613 bp in length
	Center code: BCM	41079	41178: gap of unknown length
	Web site: http://www.hgsc.bcm.tmc.edu/	41179	44036: contig of 2858 bp in length
	Contact: hgsc-help@bcm.tmc.edu	44037	44136: gap of unknown length
	----- Project Information	44137	46788: contig of 2652 bp in length
	Center project name: CH230-321022	46789	46888: gap of unknown length
	Center clone name: CH230-321022	46890	48521: contig of 1633 bp in length
	----- Summary Statistics	48522	48621: gap of unknown length
	Sequencing vector: Plasmid	48622	50480: contig of 1859 bp in length
	Chemistry: Dye-terminator Big Dye: 100% of reads	50481	50580: gap of unknown length
	Assembly program: Phrap: version 0.990329	50581	52769: contig of 2189 bp in length
	Consensus quality: 141216 bases at least Q40	52770	52869: gap of unknown length
	Consensus quality: 146793 bases at least Q30	52870	53176: contig of 2307 bp in length
	Consensus quality: 151327 bases at least Q20	53177	55276: gap of unknown length
	-----	55277	57071: contig of 1795 bp in length
	* NOTE: Estimated insert size may differ from sequence length	57072	57171: gap of unknown length
	* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).	57172	59367: contig of 2196 bp in length
	* NOTE: This is a 'working draft' sequence. It currently	59368	59467: gap of unknown length
	* consists of 66 contigs. The true order of the pieces	59468	62414: contig of 2947 bp in length
	* is not known and their order in this sequence record as	62415	62514: gap of unknown length
	* arbitrary. Gaps between the contigs are represented as	62515	65245: contig of 2731 bp in length
	* runs of N, but the exact sizes of the gaps are unknown.	65246	68118: contig of 2773 bp in length
	* This record will be updated with the finished sequence	68119	68218: gap of unknown length
	* as soon as it is available and the accession number will	68219	71828: contig of 3610 bp in length
	* be preserved.	71829	71928: gap of unknown length
		71929	75096: contig of 3168 bp in length
		75097	75196: gap of unknown length
		75197	77665: contig of 2469 bp in length
		77666	77765: gap of unknown length
		80433	80433: contig of 2668 bp in length
		80434	80533: gap of unknown length
		80534	83785: contig of 3252 bp in length
		83786	83885: gap of unknown length
		83886	86640: contig of 2755 bp in length

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

```

* 86641 86740: gap of unknown length
* 86741 88907: contig of 2167 bp in length
* 88908 89007: gap of unknown length
* 89008 91758: contig of 2751 bp in length
* 91759 91858: gap of unknown length
* 91859 95513: contig of 3655 bp in length
* 95514 95613: gap of unknown length
* 95614 97538: contig of 1925 bp in length
* 97539 97638: gap of unknown length
* 97639 101107: contig of 3469 bp in length
* 101108 101207: gap of unknown length
* 101208 105010: contig of 3803 bp in length
* 105011 105110: gap of unknown length
* 105111 108252: contig of 3141 bp in length
* 108253 108351: gap of unknown length
* 110710 110710: contig of 2359 bp in length
* 110711 110810: gap of unknown length
* 110811 113807: contig of 2997 bp in length
* 113808 113907: gap of unknown length
* 113908 116303: contig of 2396 bp in length
* 116304 116403: gap of unknown length

Query Match      0.7%; Score 22; DB 12; Length 180005;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 492 AGAATATACACTGTTTATCT 513
Db 22158 AGAATATACACTGTTTATCT 22179

RESULT 121
AC174595/c
LOCUS
DEFINITION
AC174595
AC174595.2 GI:86199977
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Bos taurus (cattle)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 187676)
Muzny, D., Adams, C., Agbai II, O., Allen, C., Albrooks, S., Archer, P.,
Arredondo, H., Bandaranaike, D., Bangura, L., Baltzan, B., Beltran, R.,
Beraducci, A., Biswal, K., Blyth, P., Bonham, H., Buhay, C., Burch, P.,
Cadore, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I.,
Chacko, J., Chahrouh, M., Chavez, D., Chen, A., Chen, G., Chen, R.,
Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyie, M., Cree, A.,
Curry, S., Dai, J., Davila, M., Davis, C., Davy-Carroll, L., De
Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H.,
Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K.,
Dziuda, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M.,
Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P.,
Fowler, G., Fu, Q., Fu, E., Garcia, A., Garcia, R., Garner, T.,
Gaskin, C., Gensch, S., Ghose, S., Gill, R., Gonzalez, D.,
Gonzalez-Garay, M., Guevara, W., Holder, B., Hoiland, W., Haeblerlein, K.,
Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P.,
Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J.,
Hines, S., Hitchens, M., Hodgson, A., Hogues, M., Hollins, B.,
Howell, L., Huang, H., Johnson, B., Johnson, R., Kalafus, K., Kelly, S.,
Jacob, L., Jiang, H., Johnson, J., Imo, K., Jackson, A., Jackson, L.,
Keys, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Lara, P.,
Leal, S., Lee, K., Lee, S., Legall, F. I., Lemon, S., Lewis, L., Li, B.,
Li, Y., Li, Z., Linnell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D.,
London, P., Lopez, J., Lorensu, L., Lozada, R., Luk, T., Madu, R.,
Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E.,
McClelland, H., McPherson, J., Mercadao, C., Metzker, M.,
Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M.,
Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Eastaugh, E.,
Nott, A., Nwaokemele, O., Obregon, M., Ochi-Okorie, C., Odeh, E.,

```

```

Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B.,
Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T.,
Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quiroz, J.,
Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S.,
Rives, C., Rodriguez, F., Rojas, A., Ruiz, S. J., Sana, M., Sanders, W.,
Santibanez, J., Santos, R., Saverly, G., Scherer, S., Shen, H., Shen, Y.,
Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R.,
Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R.,
Trejos, Z., Umani, K., Vargo, C., Verduzco, D., Villasana, D., Virk, D.,
Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J.,
Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R.,
Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L.,
Zhang, Z., Zhou, J., Weinstein, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 187676)
Worley, K.C.
Direct Submission
Submitted (19-DEC-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 187676)
Bovine Genome Sequencing Consortium
Direct Submission
Submitted (31-JAN-2006) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jan 31, 2006 this sequence version replaced gi:83722266.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FATM
Center clone name: CH240-1507
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 184871 bases at least Q40
Consensus quality: 185703 bases at least Q30
Consensus quality: 186355 bases at least Q20
Estimated insert size: 189157; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a working draft sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
3726: contig of 3726 bp in length
3727 3776: gap of 50 bp
3777 17826: contig of 14050 bp in length
17827 17876: gap of 50 bp
17877 18992: contig of 2016 bp in length
18993 19942: gap of 50 bp

```

19943 46456: contig of 26514 bp in length
46508: gap of 50 bp
46507 52968: contig of 6462 bp in length
53018: gap of 50 bp
53019 53018: contig of 2333 bp in length
55401: gap of 50 bp
57888: contig of 2487 bp in length
57938: gap of 50 bp
90149: contig of 32211 bp in length
90150 90150: gap of 50 bp
90200 94796: contig of 4597 bp in length
94797 94846: gap of 50 bp
101411: contig of 6565 bp in length
101412 101461: gap of 50 bp
102671: contig of 1210 bp in length
102721: gap of 50 bp
124959: contig of 2238 bp in length
125009: gap of 50 bp
125010 140365: contig of 15356 bp in length
140415: gap of 50 bp
140416 185481: contig of 45066 bp in length
185482 185581: gap of unknown length
185582 187676: contig of 2095 bp in length.

Location/Qualifiers
1. .187676
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-1507"
3727. .3776
/estimated_length=50
17827. .17876
/estimated_length=50
19893. .19942
/estimated_length=50
46457. .46506
/estimated_length=50
52969. .53018
/estimated_length=50
55352. .55401
/estimated_length=50
57889. .57938
/estimated_length=50
90150. .90199
/estimated_length=50
94797. .94846
/estimated_length=50
101412. .101461
/estimated_length=50
102672. .102721
/estimated_length=50
124960. .125009
/estimated_length=50
140366. .140415
/estimated_length=50
185482. .185581
/estimated_length=unknown

ORIGIN
Query Match 0.7%; Score 22; DB 12; Length 187676;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1416 TGATGTTTATGGAAGGCAGAGA 1437
|||||
DB 67939 TGATGTTTATGGAAGGCAGAGA 67918
|||||
RESULT 122
BX537287/c 195505 bp DNA linear VRT 26-FEB-2004
LOCUS Zebrafish DNA sequence from clone CH211-20608 in linkage group 19,
DEFINITION complete sequence.

ACCESSION BX537287
VERSION BX537287.8 GI:42820909
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 195505)
Barker,G.
Direct Submission
Submitted (26-FEB-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 25, 2004 this sequence version replaced gi:42592674.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/Clone-derived
zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.
Repeat names beginning 'Dr' were identified by the Recon repeat
discovery system (Zhiyong Bao and Sean Eddy, submitted), and those
beginning 'drr' were identified by Rick Waterman (Stephen Johnson
lab, WashU). For further information see
http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-20608
is from a CHORI-211 BAC library
VECTOR: pTARBAC2.1.

FEATURES
source
1. .195505
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-20608"
/clone_lib="CHORI-211"

ORIGIN
Query Match 0.7%; Score 22; DB 11; Length 195505;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3051 ATGCTCAAAATCAATGATACA 3072
|||||
DB 51574 ATGCTCAAAATCAATGATACA 51553
|||||
RESULT 123
AP000847/c 196928 bp DNA linear HTG 30-MAY-2000
LOCUS AP000847
DEFINITION Homo sapiens chromosome 11 clone RP11-773L2 map 11q24, WORKING


```

ACCESSION      AP000847
VERSION        AP000847.2
KEYWORDS       HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE         Homo sapiens (human)
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Homidae; Homo.
REFERENCE      1 (bases 1 to 196928)
               Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
               Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
               Homo sapiens 196,928 genomic DNA of 11q24
               Published Only in DataBase (1999)
REFERENCE      2 (bases 1 to 196928)
               Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
               Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
               Direct Submission
               Submitted (06-DEC-1999) Masahira Hattori, The Institute of Physical
               and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
               Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
               Japan (E-mail: hattori@gsc.riken.go.jp,
               URL: http://hgp.gsc.riken.go.jp/, Tel: 81-42-778-9923,
               Fax: 81-42-778-9924)
               On May 31, 2000 this sequence version replaced gi:6957679.
COMMENT       ----- Genome Center
               Center: RIKEN Genomic Sciences Center (GSC)
               Center code: RIKEN
               Web site: http://hgp.gsc.riken.go.jp/
               Contact: hattori@gsc.riken.go.jp
               ----- Project Information
               Center project name: HumDraft11
               Center clone name: RP11-773L2
               ----- Summary Statistics
               Sequencing vector: PCR products; 100% of reads
               Chemistry: Dye-terminator ET-amersham; 100% of reads
               Assembly program: Phrap; version 0.990329
               Consensus quality: 1797.5 bases at least Q40
               Consensus quality: 1804.3 bases at least Q30
               Consensus quality: 1917.83 bases at least Q20
               Insert size: 193928; sum-of-contigs
               Quality coverage: 4.17x in Q20 bases; sum-of-contigs
               -----
NOTE: This is a 'working draft' sequence. It currently consists of
31 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved.
1
13348 contig of 13348 bp in length
13449 gap of 100 bp
27090 contig of 11607 bp in length
38797 contig of 12107 bp in length
51004 contig of 9975 bp in length
61079 contig of 9901 bp in length
73970 contig of 9963 bp in length
82905 contig of 9963 bp in length
92906 contig of 9963 bp in length
102969 contig of 9963 bp in length
110834 contig of 9963 bp in length
110835 contig of 9963 bp in length
110935 contig of 9963 bp in length
118975 contig of 9963 bp in length
118975 contig of 9963 bp in length
126648 contig of 9963 bp in length
126648 contig of 9963 bp in length
135519 contig of 9963 bp in length
135519 contig of 9963 bp in length
141551 contig of 9963 bp in length
141552 contig of 9963 bp in length
148610 contig of 9963 bp in length
148710 contig of 9963 bp in length
154320 contig of 9963 bp in length
154320 contig of 9963 bp in length
160564 contig of 9963 bp in length
160564 contig of 9963 bp in length
167021 contig of 9963 bp in length
167021 contig of 9963 bp in length
171117 contig of 9963 bp in length
171117 contig of 9963 bp in length
173789 contig of 9963 bp in length
173789 contig of 9963 bp in length
178186 contig of 9963 bp in length
178186 contig of 9963 bp in length
180624 contig of 9963 bp in length
180624 contig of 9963 bp in length
180724 contig of 9963 bp in length
180724 contig of 9963 bp in length
182354 contig of 9963 bp in length
182354 contig of 9963 bp in length
184552 contig of 9963 bp in length
184552 contig of 9963 bp in length
184622 contig of 9963 bp in length
184622 contig of 9963 bp in length
186804 contig of 9963 bp in length
186804 contig of 9963 bp in length
186904 contig of 9963 bp in length
186904 contig of 9963 bp in length
189123 contig of 9963 bp in length
189123 contig of 9963 bp in length
191157 contig of 9963 bp in length
191157 contig of 9963 bp in length
192980 contig of 9963 bp in length
192980 contig of 9963 bp in length
193080 gap of 100 bp
192981
13348: contig of 13348 bp in length
13449: gap of 100 bp
27090: contig of 13541 bp in length
28990: gap of 100 bp
38696: contig of 11607 bp in length
38797: gap of 100 bp
50904: contig of 12107 bp in length
51003: gap of 100 bp
60978: contig of 9975 bp in length
61078: gap of 100 bp
73969: contig of 12891 bp in length
74069: gap of 100 bp
82804: contig of 8735 bp in length
82905: gap of 100 bp
92905: contig of 9901 bp in length
92905: gap of 100 bp
102868: contig of 9963 bp in length
102969: gap of 100 bp
110834: contig of 7866 bp in length
110935: gap of 100 bp
118975: contig of 7940 bp in length
118975: gap of 100 bp
126647: contig of 7673 bp in length
126647: gap of 100 bp
135518: contig of 8771 bp in length
135518: gap of 100 bp
141551: contig of 5833 bp in length
141551: gap of 100 bp
148609: contig of 7058 bp in length
148709: gap of 100 bp
154319: contig of 5610 bp in length
154319: gap of 100 bp
160563: contig of 6144 bp in length
160563: gap of 100 bp
167020: contig of 6357 bp in length
167020: gap of 100 bp
171216: contig of 4096 bp in length
171216: gap of 100 bp
173788: contig of 4472 bp in length
173788: gap of 100 bp
178185: contig of 2297 bp in length
178185: gap of 100 bp
180623: contig of 2338 bp in length
180623: gap of 100 bp
182353: contig of 1630 bp in length
182353: gap of 100 bp
184621: contig of 2168 bp in length
184621: gap of 100 bp
186803: contig of 2082 bp in length
186803: gap of 100 bp
189122: contig of 2219 bp in length
189122: gap of 100 bp
191156: contig of 1934 bp in length
191156: gap of 100 bp
192980: contig of 1724 bp in length
192980: gap of 100 bp
193080: gap of 100 bp
192981

```

FEATURES	source		1. .196928		/organism="Homo sapiens"		/mol_type="genomic DNA"		/db_xref="taxon:9606"		/chromosomes="11"		/map="11q24"		/clone="RP11-773L2"		1. .13348		/note="assembly_fragment"		13449. .26989		/note="assembly_fragment"		27090. .38696		/note="assembly_fragment"		38797. .50903		/note="assembly_fragment clone_end:SP6 vector_side:right"		51004. .60978		/note="assembly_fragment"		61079. .73969		/note="assembly_fragment clone_end:T7 vector_side:left"		74070. .82804		/note="assembly_fragment"		82905. .92805		/note="assembly_fragment"		92906. .102868		/note="assembly_fragment"		102969. .110834		/note="assembly_fragment"		110935. .118874		/note="assembly_fragment"		118975. .126647		/note="assembly_fragment"		126748. .135518		/note="assembly_fragment"		135619. .141451		/note="assembly_fragment"		141552. .148609		/note="assembly_fragment"		148710. .154319		/note="assembly_fragment"		154420. .160563		/note="assembly_fragment"		160664. .167020		/note="assembly_fragment"		167121. .171216		/note="assembly_fragment"		171317. .175788		Query Match		0.7%; Score 22; DB 12; Length 196928;		Best Local Similarity		100.0%; Pred. No. 5.8;		Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		QY 3098 ATAGAAATCTATAAAAAAAA 3119				Db 106953 ATAGAAATCTATAAAAAAAA 106932				RESULT 124		AC123934/c		LOCUS		AC123934		Mus musculus BAC clone RP23-354D22		DNA		207817 bp		linear		ROD 11-NOV-2003		DEFINITION		Mus musculus BAC clone RP23-354D22 from chromosome 17, complete sequence.		AC123934		AC123934.2		GI:27657606		VERSION		HTG.		KEYWORDS		Mus musculus (house mouse)		SOURCE		Mus musculus		ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	source		1. .196928		/organism="Homo sapiens"		/mol_type="genomic DNA"		/db_xref="taxon:9606"		/chromosomes="11"		/map="11q24"		/clone="RP11-773L2"		1. .13348		/note="assembly_fragment"		13449. .26989		/note="assembly_fragment"		27090. .38696		/note="assembly_fragment"		38797. .50903		/note="assembly_fragment clone_end:SP6 vector_side:right"		51004. .60978		/note="assembly_fragment"		61079. .73969		/note="assembly_fragment clone_end:T7 vector_side:left"		74070. .82804		/note="assembly_fragment"		82905. .92805		/note="assembly_fragment"		92906. .102868		/note="assembly_fragment"		102969. .110834		/note="assembly_fragment"		110935. .118874		/note="assembly_fragment"		118975. .126647		/note="assembly_fragment"		126748. .135518		/note="assembly_fragment"		135619. .141451		/note="assembly_fragment"		141552. .148609		/note="assembly_fragment"		148710. .154319		/note="assembly_fragment"		154420. .160563		/note="assembly_fragment"		160664. .167020		/note="assembly_fragment"		167121. .171216		/note="assembly_fragment"		171317. .175788		Query Match		0.7%; Score 22; DB 12; Length 196928;		Best Local Similarity		100.0%; Pred. No. 5.8;		Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		QY 3098 ATAGAAATCTATAAAAAAAA 3119				Db 106953 ATAGAAATCTATAAAAAAAA 106932				RESULT 124		AC123934/c		LOCUS		AC123934		Mus musculus BAC clone RP23-354D22		DNA		207817 bp		linear		ROD 11-NOV-2003		DEFINITION		Mus musculus BAC clone RP23-354D22 from chromosome 17, complete sequence.		AC123934		AC123934.2		GI:27657606		VERSION		HTG.		KEYWORDS		Mus musculus (house mouse)		SOURCE		Mus musculus		ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

Mammalia; Euthera; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus.

1 (bases 1 to 207817)

Ali, J., Haglund, K., Schatzkamer, K., Mangiapanello, L. and Haakenson, W.

The sequence of Mus musculus BAC clone RP23-354D22

Unpublished (2001)

2 (bases 1 to 207817)

Wilson, R.

Sequencing of Mus musculus

Unpublished (2001)

3 (bases 1 to 207817)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (05-JUN-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

4 (bases 1 to 207817)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (12-JAN-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

5 (bases 1 to 207817)

Wilson, R.

Direct Submission

Submitted (11-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Jan 12, 2003 this sequence version replaced gi:21326391.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu>

Contact: submissions@watson.wustl.edu

----- Summary Statistics

Center project name: M_BA0354D22

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Oseagawa and Minako Tateo in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57Bl/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC105304 and AC109612.

Location/Qualifiers

1. .207817

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="17"

/map="17"

FEATURES

source

```
repeat_region /clone="RP23-354D22"
repeat_region /clone.lib="RPCI-23"
repeat_region 935..1281
repeat_region /rpt_family="MaLR"
repeat_region 1296..1508
repeat_region /rpt_family="MaLR"
repeat_region 1536..1718
repeat_region /rpt_family="MaLR"
repeat_region 1870..2093
repeat_region /rpt_family="ERV1"
repeat_region 2833..3197
repeat_region /rpt_family="MaLR"
repeat_region 4304..4781
repeat_region /rpt_family="B4"
repeat_region 6763..6882
repeat_region /rpt_family="MIR"
repeat_region 6953..7247
repeat_region /rpt_family="B4"
repeat_region 10130..10214
repeat_region /rpt_family="B4"
repeat_region 10626..11096
repeat_region /rpt_family="RMER19B"
repeat_region 11115..11173
repeat_region /rpt_family="RMER19A"
repeat_region 11319..11557
repeat_region /rpt_family="MaLR"
misc_feature 13777..13950
repeat_region /note="Sequence derived from PCR product of project DNA."
repeat_region 14130..14374
repeat_region /rpt_family="MaLR"
repeat_region 14380..14603
repeat_region /rpt_family="MaLR"
repeat_region 14769..14965
repeat_region /rpt_family="MaLR"
repeat_region 15019..15835
repeat_region /rpt_family="MaLR"
repeat_region 15903..16231
repeat_region /rpt_family="MaLR"
repeat_region 16235..16601
repeat_region /rpt_family="MaLR"
repeat_region 19268..19438
repeat_region /rpt_family="B2"
repeat_region 19847..20011
repeat_region /rpt_family="L1"
repeat_region 22856..23011
repeat_region /rpt_family="B4"
repeat_region 24412..24612
repeat_region /rpt_family="RMER10B"
repeat_region 24740..24780
repeat_region /rpt_family="RMER17C"
repeat_region 24975..25195
repeat_region /rpt_family="B4"
repeat_region 26153..26442
repeat_region /rpt_family="L1"
repeat_region 26878..27099
repeat_region /rpt_family="MER1_type"
repeat_region 28861..28997
repeat_region /rpt_family="Alu"
repeat_region 28998..29146
repeat_region /rpt_family="Alu"
repeat_region 30673..30827
repeat_region /rpt_family="ERVK"
repeat_region 31518..31628
repeat_region /rpt_family="MaLR"
repeat_region 34256..34311
repeat_region /rpt_family="U3"
repeat_region 35039..35245
repeat_region /rpt_family="L1"
repeat_region 35249..35405
repeat_region /rpt_family="L1"
repeat_region 35646..35711
repeat_region /rpt_family="L1"
repeat_region 36070..36505
```

```
repeat_region /rpt_family="MaLR"
repeat_region 36798..37017
repeat_region /rpt_family="B4"
repeat_region 37997..38691
repeat_region /rpt_family="L1"
repeat_region 40345..40407
repeat_region /rpt_family="ERV1"
repeat_region 41048..41202
repeat_region /rpt_family="L1"
repeat_region 41899..42336
repeat_region /rpt_family="L1"
repeat_region 42787..42842
repeat_region /rpt_family="ID"
misc_feature 42949..43062
repeat_region 43183..43298
repeat_region /rpt_family="L1"
repeat_region 43745..43854
repeat_region /rpt_family="MIR"
repeat_region 44932..44983
repeat_region /rpt_family="ERV1"
repeat_region 44963..45013
repeat_region /rpt_family="ERV1"
repeat_region 45084..45134
repeat_region /rpt_family="ERV1"
repeat_region 45191..45240
repeat_region /rpt_family="ERV1"
repeat_region 45249..45315
repeat_region /rpt_family="L1"
repeat_region 45321..45375
repeat_region /rpt_family="ERV1"
repeat_region 45355..45406
repeat_region /rpt_family="ERV1"
repeat_region 47156..47391
repeat_region /rpt_family="L1"
repeat_region 49636..49840
repeat_region /rpt_family="B4"
repeat_region 51112..51333
repeat_region /rpt_family="MER1_type"
repeat_region 55474..55547
repeat_region /rpt_family="Alu"
repeat_region 57783..57978
repeat_region /rpt_family="B2"
repeat_region 58531..58624
repeat_region /rpt_family="B2"

Query Match 0.7%; Score 22; DB 6; Length 207817;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2064 AGAAATTTTCCTTTGAAAGT 2085
Db 83187 AGAAATTTTCCTTTGAAAGT 83166

RESULT 125
AC151257/c
LOCUS AC151257
DEFINITION Bos taurus clone CH240-342N10, WORKING DRAFT SEQUENCE, 17 unordered
pieces.
AC151257
VERSION AC151257.3 GI:68265941
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Bos taurus (cattle)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 210651)
Muzny,D.Marte., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Ayoyagi,M., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
```

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FBHW
Center clone name: CH240-342N10
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 201631 bases at least Q40
Consensus quality: 203740 bases at least Q30
Consensus quality: 205693 bases at least Q20
Estimated insert size: 210305; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2052: contig of 2052 bp in length
3315: gap of 1263 bp
6310: contig of 2995 bp in length
6410: gap of unknown length
12869: contig of 6459 bp in length
12919: gap of 50 bp
35273: contig of 22354 bp in length
35323: gap of 50 bp
60389: contig of 25066 bp in length
60439: gap of 50 bp
74633: contig of 14194 bp in length
74683: gap of 50 bp
101582: contig of 26899 bp in length
101632: gap of 50 bp
128266: contig of 26634 bp in length
128316: gap of 50 bp
143654: contig of 15338 bp in length
143704: gap of 50 bp
190712: contig of 47008 bp in length
190762: gap of 50 bp
190763: contig of 7558 bp in length
198320: contig of 50 bp
198371: contig of 4686 bp in length
203056: contig of 1525 bp in length
203106: gap of 50 bp
204631: contig of 1478 bp in length
204681: gap of 50 bp
206159: contig of 1374 bp in length
206259: gap of unknown length
207633: contig of 1374 bp in length
207733: gap of unknown length
209520: contig of 1787 bp in length
209521: gap of unknown length
210651: contig of 1031 bp in length.
----- Location/Qualifiers
1. 210651
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-342N10"
2053. .3315
/estimated_length=1263
6311. .6410
/estimated_length=unknown
12870. .12919
/estimated_length=50
35274. .35323
/estimated_length=50
60390. .60439

FEATURES
source
gap
gap
gap
gap
gap
----- Genome Center

```

/estimated_length=50
74634 .74683
/estimated_length=50
101583 .101632
/estimated_length=50
128267 .128316
/estimated_length=50
143655 .143704
/estimated_length=50
190713 .190762
/estimated_length=50
198321 .198370
/estimated_length=50
203057 .203106
/estimated_length=50
204632 .204681
/estimated_length=50
206160 .206259
/estimated_length=unknown
207634 .207733
/estimated_length=unknown
209521 .209620
/estimated_length=unknown

ORIGIN
Query Match      0.7%  Score 22;  DB 12;  Length 210651;
Best Local Similarity 100.0%;  Pred. No. 5.8;
Matches 22;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY 1506 AAATATCCATGACATCTTTCAT 1527
Db 53204 AAATATCCATGACATCTTTCAT 53183

RESULT 126
AC167594      210788 bp      DNA      linear      HTG 23-JAN-2006
Bos taurus clone CH240-137M9, *** SEQUENCING IN PROGRESS ***, 25
unordered pieces.
AC167594
AC167594.2 GI:85665454
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Bos taurus (cattle)
Bos taurus

Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 210788)
Muzny, D., Adams, C., Agbai II, O., Allen, C., Alsbrooks, S., Archer, P.,
Arredondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R.,
Beraducci, A., Biswal, K., Blyth, P., Bonham, H., Buhay, C., Burch, P.,
Cadoree, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I.,
Chacko, J., Chahrouh, M., Chavez, D., Chen, K., Chen, G., Chen, R.,
Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A.,
Curry, S., Dai, W., Davila, M. L., Davis, C., Davy-Carroll, L., De
Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H.,
Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K.,
Dziuda, D., Egan, A., Escoto, M., Espinosa, V., Eugene, C., Fa, M.,
Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P.,
Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T.,
Gaskin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D.,
Gonzalez-Garay, M., Guevara, W., Holder, M., Hsaland, W., Haeblerlen, K.,
Hall, B., Hamid, H., Hamilton, K., Harber, B., Harris, R., Havlak, P.,
Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J.,
Hines, S., Hitchens, M., Hodgson, A., Hogue, M., Hollins, B.,
Howell, L. T., Hulyk, S., Hume, J., Imo, K., Jackson, A., Jackson, L.,
Jacob, L., Jiang, K., Johnson, B., Johnson, R., Kalatous, K., Kelly, S.,
Key, T., Khan, Z., King, L., Kovar, C., Kowis, C., Lard, F.,
Lea, S., Lee, K., Lee, S., LeGall, F. I., Lemon, S., Lewis, L., Li, B.,
Li, X., Li, Z., Limell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D.,
London, F., Lopez, J., Lorensuewa, L., Lozano, R., Luk, T., Madu, R.,
Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E.,
McClelland, H., McPherson, J., Mercadao, C., Metzker, M.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
CONSTRM
JOURNAL
TITLE
JOURNAL
COMMENT

Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M.,
Murray, D., Nazarith, L., Ngo, D., Nguyen, N., Norwig-Estaugh, E.,
Nott, A., Nwaokemele, O., Obregon, M., Ochi-Okorie, C., Odeh, E.,
Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B.,
Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T.,
Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quiroz, J.,
Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S.,
Rives, C., Rodriguez, F., Rojas, A., Ruiz, S. J., Sana, M., Sanders, W.,
Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, H., Shen, Y.,
Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R.,
Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R.,
Trejos, Z., Usmani, K., Varco, C., Verduzco, D., Villaseña, D., Virk, D.,
Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J.,
Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R.,
Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L.,
Zhang, Z., Zhou, J., Weinstein, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 210788)
Worley, K.C.
Direct Submission
Submitted (01-SEP-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 210788)
Bovine Genome Sequencing Consortium
Direct Submission
Submitted (23-JAN-2006) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jan 23, 2006 this sequence version replaced gi:74039134.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the ends of the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc.help@bcm.tmc.edu
----- Project Information
Center Project name: FLPI
Center Clone name: CH240-137M9
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 199274 bases at least Q40
Consensus quality: 202130 bases at least Q30
Consensus quality: 204587 bases at least Q20
Estimated insert size: 210018; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 23387: contig of 23387 bp in length
* 23388 23437: gap of 50 bp
* 23438 36183: contig of 12746 bp in length

```



```

Paternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puzos, M., Quiroz, J., Rachin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rokey, T., Rojao, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
Snedik, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, I., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
Valas, R., Vera, V., Villaseña, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczky, R., Wood, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 214864)
Worley, K.C.
Direct Submission
Submitted (31-MAY-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 214864)
Cow Genome Sequencing Consortium.
Direct Submission
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FHNA
Center clone name: CH240-124D18
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 189868 bases at least Q40
Consensus quality: 191139 bases at least Q30
Consensus quality: 192213 bases at least Q20
Estimated insert size: 194905; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 7644: contig of 7644 bp in length
* 7694: gap of 50 bp
* 7695 33036: contig of 25342 bp in length

```

```

33037 33086: gap of 50 bp
33087 43580: contig of 10494 bp in length
33088 43580: gap of 50 bp
43581 45246: contig of 1616 bp in length
45247 53945: gap of 8699 bp
53946 56598: contig of 2653 bp in length
56599 56648: gap of 50 bp
56649 68032: contig of 11384 bp in length
68033 87389: contig of 50 bp
87390 87439: contig of 19307 bp in length
87440 100942: contig of 13503 bp in length
100943 102144: contig of 1152 bp in length
102145 112702: gap of 10558 bp
112703 135129: contig of 22427 bp in length
135130 135325: gap of 196 bp
135326 144585: contig of 9260 bp in length
144586 14635: gap of 50 bp
14636 148235: contig of 3600 bp in length
148236 183472: gap of 237 bp
183473 183472: contig of 36873 bp in length
183474 183472: gap of 50 bp
183475 195071: contig of 9676 bp in length
195072 195121: gap of 50 bp
195122 197679: contig of 2558 bp in length
197680 197729: gap of 50 bp
202425 202425: contig of 4696 bp in length
202426 202827: gap of 402 bp
202828 213656: contig of 10829 bp in length
213657 213756: gap of unknown length
213757 214864: contig of 1108 bp in length.
FEATURES
source
1..214864
location/Qualifiers
organism="Bos taurus"
mol_type="genomic DNA"
db_xref="taxon:9913"
clone="CH240-124D18"
7645..7694
/estimated_length=50
33037..33086
/estimated_length=50
43581..43630
/estimated_length=50
45247..53945
/estimated_length=8699
56599..56648
/estimated_length=50
68033..68082
/estimated_length=50
87390..87439
/estimated_length=50
100943..100992
/estimated_length=50
102145..112702
/estimated_length=10558
135130..135325
/estimated_length=196
144586..144635
/estimated_length=50
148236..148472
/estimated_length=237
183473..185395
/estimated_length=50
195072..195121
/estimated_length=50
197680..197729
/estimated_length=50
202426..202827
/estimated_length=402
213657..213756
/estimated_length=unknown
ORIGIN

```


Query Match 0.7%; Score 22; DB 12; Length 214864;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1416 TGAATGTTATGGAAGCAGAGA 1437
 |||||
 Db 158460 TGAATGTTATGGAAGCAGAGA 158439

RESULT 128
 AC156845 222380 bp DNA linear HTG 01-JUL-2005
 LOCUS Bos taurus clone CH240-65E3, *** SEQUENCING IN PROGRESS ***, 23
 DEFINITION unordered pieces.

ACCESSION AC156845.2 GI:68303004
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 KEYWORDS Bos taurus (cattle)
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.
 1 (bases 1 to 222380)
 Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gregeorge, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensuhewa, L., Louised, H., Lozado, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Mlosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwaokeleneh, O., Okwuonu, G., Olarnpusagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
 Villas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wlarczyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zuo, D., Smith, H. O.,
 Niederhauser, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstock, G. and Gibbs, R. A.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Direct Submission
 Unpublished
 2 (bases 1 to 222380)
 Worley, K. C.
 Direct Submission
 Submitted (05-FEB-2005) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 222380)
 Cow Genome Sequencing Consortium.
 Direct Submission
 Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jun 29, 2005 this sequence version replaced gi:58652202.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: FDTG
 Center clone name: CH240-65E3
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 207604 bases at least Q40
 Consensus quality: 210552 bases at least Q30
 Consensus quality: 213132 bases at least Q20
 Estimated insert size: 214079; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 23 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1539: contig of 1539 bp in length
 1540 1871: gap of 332 bp
 1872 7640: contig of 5769 bp in length
 7641 7874: gap of 234 bp
 7875 23979: contig of 16105 bp in length
 24029: gap of 50 bp
 24030 29654: contig of 5625 bp in length
 29655 29865: gap of 211 bp
 29866 34956: contig of 5091 bp in length
 34957 35006: gap of 50 bp
 35007 41109: contig of 6103 bp in length
 41110 41688: gap of 579 bp
 41689 48375: contig of 6687 bp in length
 48376 48425: gap of 50 bp
 48426 84184: contig of 35759 bp in length
 84185 84234: gap of 50 bp
 84235 88842: contig of 4608 bp in length
 88843 88892: gap of 50 bp
 88893 98098: contig of 9206 bp in length
 98099 98148: gap of 50 bp

```

* 98149 101062: contig of 2914 bp in length
* 101063 102005: gap of 943 bp
* 102006 181090: contig of 79085 bp in length
* 181091 182305: gap of 1215 bp
* 182306 184195: contig of 1890 bp in length
* 184196 185630: gap of 1885 bp
* 185631 187680: contig of 1550 bp in length
* 187681 203118: gap of 50 bp
* 203119 203668: contig of 14638 bp in length
* 203669 205027: gap of 50 bp
* 205028 205077: contig of 2659 bp in length
* 205078 209276: gap of 50 bp
* 209277 209326: contig of 4199 bp in length
* 209327 215614: gap of 50 bp
* 215615 215714: contig of 6288 bp in length
* 215715 216801: gap of unknown length
* 216802 216901: contig of 1087 bp in length
* 216902 217942: gap of unknown length
* 217943 218042: contig of 1041 bp in length
* 218043 219264: gap of unknown length
* 219265 219364: contig of 1222 bp in length
* 219365 221063: gap of unknown length
* 221064 221163: contig of 1699 bp in length
* 221164 222380: gap of unknown length
* 222380 222380: contig of 1217 bp in length.
Location/Qualifiers
1. .222380
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-65E3"
1540. .1871
/estimated_length=332
7641. .7874
/estimated_length=234
23980.240029
/estimated_length=50
29655. .29665
/estimated_length=211
34957. .35006
/estimated_length=50
4110. .41688
/estimated_length=579
48376. .48425
/estimated_length=50
84185. .84234
/estimated_length=50
88843. .88892
/estimated_length=50
98099. .98148
/estimated_length=50
101063. .102005
/estimated_length=943
181091. .182305
/estimated_length=1215
184196. .186080
/estimated_length=1885
187631. .187680
/estimated_length=50
202319. .202368
/estimated_length=50
205028. .205077
/estimated_length=50
209277. .209326
/estimated_length=50
215615. .215714

```

FEATURES

```

source
1. .222380
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="C57BL/6J"
1540. .1871
/estimated_length=332
7641. .7874
/estimated_length=234
23980.240029
/estimated_length=50
29655. .29665
/estimated_length=211
34957. .35006
/estimated_length=50
4110. .41688
/estimated_length=579
48376. .48425
/estimated_length=50
84185. .84234
/estimated_length=50
88843. .88892
/estimated_length=50
98099. .98148
/estimated_length=50
101063. .102005
/estimated_length=943
181091. .182305
/estimated_length=1215
184196. .186080
/estimated_length=1885
187631. .187680
/estimated_length=50
202319. .202368
/estimated_length=50
205028. .205077
/estimated_length=50
209277. .209326
/estimated_length=50
215615. .215714

```

Query Match 0.7%; Score 22; DB 12; Length 222380;
Best Local Similarity 100.0%; Pred.No.5.8; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1424 TCGAAGGCGAGACTCATTGA 1445
|||||

Db	211020	TCGAAGGCGAGACTCATTGA	211041
RESULT 129			
AC105304		229816 bp	DNA linear
LOCUS			ROD 16-SEP-2003
DEFINITION			Mus musculus strain C57BL/6J clone rp23-71k5 map 17, complete sequence.
ACCESSION	AC105304	GI:25815339	
VERSION	AC105304.24		
KEYWORDS	HTG.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			Eukaryota; Chordata; Metazoa; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.
AUTHORS			1 (bases 1 to 229816) Deschamps, S., Li, Y., Hu, X., Brathwaite, M., Waeltz, P., Nagaraja, R. and Roe, B.A.
TITLE			Mus musculus BAC Clone rp23-71k5
REFERENCE			2 (bases 1 to 229816) Hu, X. and Roe, B.A.
AUTHORS			Direct Submission
TITLE			Submitted (29-DEC-2001) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
JOURNAL			3 (bases 1 to 229816) Deschamps, S., Li, Y., Hu, X. and Roe, B.A.
REFERENCE			Direct Submission
AUTHORS			Submitted (23-NOV-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
JOURNAL			4 (bases 1 to 229816) Deschamps, S., Li, Y., Hu, X. and Roe, B.A.
REFERENCE			Direct Submission
AUTHORS			Submitted (28-NOV-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
JOURNAL			5 (bases 1 to 229816) Deschamps, S., Li, Y., Hu, X., Brathwaite, M., Waeltz, P., Nagaraja, R. and Roe, B.A.
REFERENCE			Direct Submission
AUTHORS			Submitted (21-MAR-2003) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
JOURNAL			6 (bases 1 to 229816) Deschamps, S., Li, Y., Hu, X., Brathwaite, M., Waeltz, P., Nagaraja, R. and Roe, B.A.
REFERENCE			Direct Submission
AUTHORS			Submitted (16-SEP-2003) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
JOURNAL			On Nov 28, 2002 this sequence version replaced gi:25141046.
COMMENT			----- Genomic Center Center: Department Of Chemistry And Biochemistry The University Of Oklahoma Center code:UOKNOR
FEATURES			Location/Qualifiers 1. .229816 /organism="Mus musculus" /mol_type="genomic DNA" /strain="C57BL/6J" /db_xref="taxon:10090" /map="17" /clone="rp23-71k5" /clone_lib="RPC1 - 23 Female (C57BL/6J) Mouse BAC Library"
ORIGIN			
Query Match			0.7%; Score 22; DB 6; Length 229816;
Best Local Similarity			100.0%; Pred.No.5.8;
Matches			22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

TITLE	Direct Submission
JOURNAL	Submitted (14-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 231295)
AUTHORS	Rat Genome Sequencing Consortium.
TITLE	Direct Submission
JOURNAL	Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On May 10, 2003 this sequence version replaced gi:24635699. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
	----- Genome Center
	Center: Baylor College of Medicine
	Center code: BCM
	Web site: http://www.hgsc.bcm.tmc.edu/
	Contact: hgsc-help@bcm.tmc.edu
	----- Project Information
	Center project name: GGOZ
	Center clone name: CH230-19G10
	----- Summary Statistics
	Assembly program: Atlas 3.0;
	Consensus quality: 223363 bases at least Q40
	Consensus quality: 225310 bases at least Q30
	Consensus quality: 226539 bases at least Q20
	Estimated insert size: 231185; sum-of-contigs estimation
	Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

	* NOTE: Estimated insert size may differ from sequence length
	* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
	* NOTE: This is a 'working draft' sequence. It currently
	* consists of 3 contigs. The true order of the pieces
	* is not known and their order in this sequence record is
	* arbitrary. Gaps between the contigs are represented as
	* runs of N, but the exact sizes of the gaps are unknown.
	* This record will be updated with the finished sequence
	* as soon as it is available and the accession number will
	* be preserved.
	* 1 228572: contig of 228572 bp in length
	* 228573 228672: gap of unknown length
	* 228673 230105: contig of 1433 bp in length
	* 230106 230205: gap of unknown length
	* 230206 231295: contig of 1090 bp in length.
FEATURES	Location/Qualifiers
source	1..231295
	/organism="Rattus norvegicus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10116"
	/clone="CH230-19G10"
	1421..2288
	/note="clone boundary
	clone_end:Sp6
	site:EcoRI
	end sequence:BH269139"
	complement(226435..227317)
	/note="clone boundary
	clone_end:T7
	site:EcoRI
	end sequence:BH269138"
	228573..228672
	/estimated_length=unknown
	230106..230205
misc_feature	
misc_feature	
gap	
gap	
TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 231295)
AUTHORS	Worley, K.C.

ORIGIN	/estimated_length=unknown	
Query Match	0.7%; Score 22; DB 12; Length 231295;	
Best local similarity	100.0%; Pred. No. 5.8;	
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY 492	AGAAATACACTGTTTATTCT 513	
DB 43891	AGAAATACACTGTTTATTCT 43870	
RESULT 131		
LOCUS		
CT030007		
DEFINITION	Danio rerio chromosome 19 clone CH73-129D20, *** SEQUENCING IN	
ACCESSION	CT030007	
VERSION	CT030007.1	
KEYWORDS	GI:76780690	
SOURCE	HTG; HTGS PHASE1.	
ORGANISM	Danio rerio (zebrafish)	
REFERENCE	Danio rerio	
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
TITLE	Actinopterygii; Neopterygii; Teleostei; Ostariophysi;	
JOURNAL	Cypriniformes; Cyprinidae; Danio.	
COMMENT	1 (bases 1 to 258530)	
	Sims,S.	
	Direct Submission	
	Submitted (30-SEP-2005) Wellcome Trust Sanger Institute, Hinxton,	
	Cambridgeshire, CB10 1SA, UK. E-mail enquiries:	
	zfish-help@sanger.ac.uk Clone requests:	
	http://www.sanger.ac.uk/Projects/D_rerio/faqs.shtml#dataeight	
	----- Genome Center	
	Center: Wellcome Trust Sanger Institute	
	Center code: SC	
	Web site: http://www.sanger.ac.uk	
	Contact: zfish-help@sanger.ac.uk	
	----- Project Information	
	Center project name: zh129D20	
	----- Summary Statistics	
	Assembly program: XGAP4; version 4.5	
	Chemistry: Dye-terminator; 100% of reads	
	Consensus quality: 251718 bases at least Q40	
	Consensus quality: 253600 bases at least Q30	
	Consensus quality: 254798 bases at least Q20	
	Insert size: 256730; sum-of-contigs	
	Insert size: 152668; 7.1% error; agarose-fp	
	Quality coverage: 6.04x in Q20 bases; sum-of-contigs Quality	
	coverage: 10.28x in Q20 bases; agarose-fp	

	* NOTE: This is a 'working draft' sequence. It currently	
	* consists of 19 contigs. The true order of the pieces	
	* is not known and their order in this sequence record is	
	* arbitrary. Gaps between the contigs are represented as	
	* runs of N, but the exact sizes of the gaps are unknown.	
	* This record will be updated with the finished sequence.	
	* as soon as it is available and the accession number will	
	* be preserved.	

	1 3719: contig of 3619 bp in length	
	3620 3719: gap of 100 bp	
	3720 23103: contig of 19384 bp in length	
	23104 23203: gap of 100 bp	
	23204 31532: contig of 8329 bp in length	
	31533 31632: gap of 100 bp	
	31633 45612: contig of 13980 bp in length	
	45613 45712: gap of 100 bp	
	45713 52360: contig of 6648 bp in length	
	52361 52461: gap of 100 bp	
	52461 63407: contig of 10946 bp in length	
	63407 71024: contig of 100 bp	
	71025 71124: gap of 100 bp	
	71125 75011: contig of 3887 bp in length	
	75012 75011: contig of 9905 bp in length	
	75013 75012: gap of 100 bp	
	75014 75013: contig of 5094 bp in length	
	75015 75014: gap of 100 bp	
	75016 75015: contig of 7563 bp in length	
	75017 75016: gap of 100 bp	
	75018 75017: contig of 2132 bp in length	
	75019 75018: gap of 100 bp	
	75020 75019: contig of 5643 bp in length	
	75021 75020: gap of 100 bp	
	75022 75021: contig of 3807 bp in length	
	75023 75022: gap of 100 bp	
	75024 75023: contig of 62867 bp in length	
	75025 75024: gap of 100 bp	
	75026 75025: contig of 56406 bp in length	
	75027 75026: gap of 100 bp	
	75028 75027: contig of 7487 bp in length	
	75029 75028: gap of 100 bp	
	75030 75029: contig of 10040 bp in length	
	75031 75030: gap of 100 bp	
	75032 75031: contig of 11475 bp in length	
	75033 75032: gap of 100 bp	
	75034 75033: contig of 11475 bp in length	
	75035 75034: gap of 100 bp	
	75036 75035: contig of 11475 bp in length	
	75037 75036: gap of 100 bp	
	75038 75037: contig of 11475 bp in length	
	75039 75038: gap of 100 bp	
	75040 75039: contig of 11475 bp in length	
	75041 75040: gap of 100 bp	
	75042 75041: contig of 11475 bp in length	
	75043 75042: gap of 100 bp	
	75044 75043: contig of 11475 bp in length	
	75045 75044: gap of 100 bp	
	75046 75045: contig of 11475 bp in length	
	75047 75046: gap of 100 bp	
	75048 75047: contig of 11475 bp in length	
	75049 75048: gap of 100 bp	
	75050 75049: contig of 11475 bp in length	
	75051 75050: gap of 100 bp	
	75052 75051: contig of 11475 bp in length	
	75053 75052: gap of 100 bp	
	75054 75053: contig of 11475 bp in length	
	75055 75054: gap of 100 bp	
	75056 75055: contig of 11475 bp in length	
	75057 75056: gap of 100 bp	
	75058 75057: contig of 11475 bp in length	
	75059 75058: gap of 100 bp	
	75060 75059: contig of 11475 bp in length	
	75061 75060: gap of 100 bp	
	75062 75061: contig of 11475 bp in length	
	75063 75062: gap of 100 bp	


```

/db_xref="taxon:10116"
/clone="CH230-204F23"
1..1665
/notes="ags end extension
clone_end:sp6"
3110..4758
/notes="wgs_end extension
clone_end:sp6"
5478..6171
/notes="clone boundary
clone_end:sp6
site:EcoRI
end_sequence:RWBjW36TV"
complement(288029..288819)
/notes="clone boundary
clone_end:T7
site:EcoRI
end_sequence:RWBjW36TJ"

ORIGIN
Query Match 0.7%; Score 22; DB 12; Length 294090;
Best Local Similarity 100.0%; Pred. No. 5.8; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0;

Qy 1853 TGACAGGCTGACTGACCGTGG 1874
|||||
Db 268679 TGACAGGCTGACTGACCGTGG 268700

RESULT 133
CQ690820
LOCUS 282 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 35746 from Patent WO02070737.
ACCESSION CQ690820
VERSION CQ690820.1 GI:42228399
KEYWORDS Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1
AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 35746 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source
1..282
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 20; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;

Qy 2422 GGCTGGTCTCTATGGAGGATAC 2442
|||||
Db 233 GGCTGGTCTCTATGGAGGATAC 253

RESULT 134
AJ865196
LOCUS 400 bp DNA linear PLN 03-MAY-2005
DEFINITION Cocos nucifera microsatellite DNA, clone CnCir 94.
ACCESSION AJ865196
VERSION AJ865196.1 GI:57898801
KEYWORDS microsatellite; repetitive DNA.
SOURCE Cocos nucifera (coconut palm)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Arecales; Arecoidae;
Cocoseae; Butiinae; Cocos.

1
Lebrun,P.
AUTHORS
TITLE CRDXRIT: A genetic coconut map
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 400)
AUTHORS Lebrun,P.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-2004) Lebrun P., CIRAD, TA 80/03 Avenue du Val de
Montferand, 34398 Montpellier Cedex 5, FRANCE
FEATURES
source
1..400
/organism="Cocos nucifera"
/mol_type="genomic DNA"
/db_xref="taxon:13894"
/clone="CnCir_94"
/tissue_type="leaf"
/country="Malaysia"
86..101
/notes="Forward Primer: CCTCCCACTCCGTATT"
168..201
/notes="microsatellite"
/rpt_type="RANDEM"
/rpt_unit_range=168..201
complement(369..383)
/notes="Reverse Primer: GGCACGGCAAGTAG"

ORIGIN
Query Match 0.7%; Score 21; DB 4; Length 400;
Best Local Similarity 100.0%; Pred. No. 20; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;

Qy 3100 AAGAAATTACTAAAAA 3120
|||||
Db 29 AAGAAATTACTAAAAA 49

RESULT 135
AR280920
LOCUS 535 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 428 from patent US 6518237.
ACCESSION AR280920
VERSION AR280920.1 GI:29716390
KEYWORDS Unknown.
ORGANISM
Unclassified.
REFERENCE 1 (bases 1 to 535)
AUTHORS Yugu,J Dillon,D.C.; Mitcham,J.L. and Xu,J.
TITLE Compositions for treatment and diagnosis of breast cancer and
methods for their use
JOURNAL Patent: US 6518237-A 428 11-FEB-2003;
Corixa Corporation; Seattle, WA
FEATURES
source
1..535
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;

Qy 2422 GGCTGGTCTCTATGGAGGATAC 2442
|||||
Db 306 GGCTGGTCTCTATGGAGGATAC 326

RESULT 136
AR283416
LOCUS 535 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 428 from patent US 6528054.
ACCESSION AR283416

```

JOURNAL Patent: US 6586572-A 428 01-JUL-2003;
Corixa Corporation; Seattle, WA
FEATURES
source
1. .535
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2422 GGCTGGTCTTATGGAGGATAC 2442
|||||
Db 306 GGCTGGTCTTATGGAGGATAC 326

RESULT 139
AR453965
LOCUS AR453965 535 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 428 from patent US 6680197.
ACCESSION AR453965
VERSION AR453965.1 GI:42686755
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 535)
AUTHORS Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J., Harlocker, S.L.,
Hepler, W.T. and Henderson, R.A.
TITLE Compositions and methods for the therapy and diagnosis of breast
cancer
JOURNAL Patent: US 6680197-A 428 20-JAN-2004;
Corixa Corporation; Seattle, WA
FEATURES
source
1. .535
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2422 GGCTGGTCTTATGGAGGATAC 2442
|||||
Db 306 GGCTGGTCTTATGGAGGATAC 326

RESULT 140
AR561553
LOCUS AR561553 535 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 428 from patent US 6756477.
ACCESSION AR561553
VERSION AR561553.1 GI:53974461
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 535)
AUTHORS Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J. and Harlocker, S.L.
TITLE Compositions and methods for the therapy and diagnosis of breast
cancer
JOURNAL Patent: US 6756477-A 428 29-JUN-2004;
Corixa Corporation; Seattle, WA
FEATURES
source
1. .535
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2422 GGCTGGTCTTATGGAGGATAC 2442
|||||
Db 306 GGCTGGTCTTATGGAGGATAC 326

RESULT 138
AR351385
LOCUS AR351385 535 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 428 from patent US 6586572.
ACCESSION AR351385
VERSION AR351385.1 GI:33753064
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 535)
AUTHORS Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J., Harlocker, S.L. and
Hepler, W.T.
TITLE Compositions and methods for the therapy and diagnosis of breast
cancer


```

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2422 GGCTGGTCCTATGGAGGATAC 2442
Db 306 GGCTGGTCCTATGGAGGATAC 326

RESULT 141
AR632008
LOCUS AR632008
DEFINITION Sequence 428 from patent US 6844325.
ACCESSION AR632008
VERSION AR632008.1 GI:59772953
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 535)
AUTHORS Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J. and Harlocker, S.L.
TITLE Compositions and methods for the treatment and diagnosis of breast cancer and methods for their use
JOURNAL Patent: US 6844325-A 428 18-JAN-2005;
Corixa Corporation; Seattle, WA
FEATURES
source
1..535
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2422 GGCTGGTCCTATGGAGGATAC 2442
Db 306 GGCTGGTCCTATGGAGGATAC 326

RESULT 142
AR760254
LOCUS AR760254
DEFINITION Sequence 428 from patent US 6958361.
ACCESSION AR760254
VERSION AR760254.1 GI:83327054
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 535)
AUTHORS Houghton, R.L., Sleath, P.R. and Persing, D.H.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer.
JOURNAL Patent: US 6958361-A 428 25-OCT-2005;
Corixa Corporation; Seattle, WA
FEATURES
source
1..535
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2422 GGCTGGTCCTATGGAGGATAC 2442
Db 306 GGCTGGTCCTATGGAGGATAC 326

RESULT 143
AR773397
LOCUS AR773397
DEFINITION Sequence 428 from patent US 6969518.
ACCESSION AR773397
VERSION AR773397.1 GI:83349927
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 535)
AUTHORS Houghton, R.L., Sleath, P.R. and Persing, D.H.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer.
JOURNAL Patent: US 6969518-A 428 29-NOV-2005;
Corixa Corporation; Seattle, WA
FEATURES
source
1..535
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2422 GGCTGGTCCTATGGAGGATAC 2442
Db 306 GGCTGGTCCTATGGAGGATAC 326

RESULT 144
AX303108
LOCUS AX303108
DEFINITION Sequence 428 from Patent WO0179286.
ACCESSION AX303108
VERSION AX303108.1 GI:17383609
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1
AUTHORS Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J., Harlocker, S.L. and Hepler, W.T.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer.
JOURNAL Patent: WO 0179286-A 428 25-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
source
1..535
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2422 GGCTGGTCCTATGGAGGATAC 2442
Db 306 GGCTGGTCCTATGGAGGATAC 326

RESULT 145
BV343295/c
LOCUS BV343295/c
DEFINITION Tagged site.
ACCESSION BV343295
VERSION BV343295.1 GI:57543598
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1
AUTHORS Canis familiaris (dog)
TITLE Canis familiaris
JOURNAL Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

ACCESSION AR773397
VERSION AR773397.1 GI:83349927
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 535)
AUTHORS Houghton, R.L., Sleath, P.R. and Persing, D.H.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer.
JOURNAL Patent: US 6969518-A 428 29-NOV-2005;
Corixa Corporation; Seattle, WA
FEATURES
source
1..535
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2422 GGCTGGTCCTATGGAGGATAC 2442
Db 306 GGCTGGTCCTATGGAGGATAC 326

RESULT 144
AX303108
LOCUS AX303108
DEFINITION Sequence 428 from Patent WO0179286.
ACCESSION AX303108
VERSION AX303108.1 GI:17383609
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1
AUTHORS Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J., Harlocker, S.L. and Hepler, W.T.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer.
JOURNAL Patent: WO 0179286-A 428 25-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
source
1..535
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2422 GGCTGGTCCTATGGAGGATAC 2442
Db 306 GGCTGGTCCTATGGAGGATAC 326

RESULT 145
BV343295/c
LOCUS BV343295/c
DEFINITION Tagged site.
ACCESSION BV343295
VERSION BV343295.1 GI:57543598
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1
AUTHORS Canis familiaris (dog)
TITLE Canis familiaris
JOURNAL Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae; Canis.

1 (bases 1 to 588)
Lindblad-Toh, K.
The genome sequence of Canis familiaris
Unpublished (2004)

Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kerli@genome.wi.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 588

Protocol:
WGS-discovery (WGS):
Paired-end low-coverage whole genome shotgun reads were generated
from 9 breeds
(German Shepherd, Rottweiler, Bedlington Terrier, Beagle, Labrador
Retriever, English
Shepherd, Italian Greyhound, Alaskan Malamute and the Portuguese
Water Dog -100,000 each)
and five other canids (Chinese, Alaskan, Indian and Spanish Gray
Wolf as well as the
Californian Coyote).

The WGS reads were placed uniquely on the CanFam1.0 boxer assembly
and SNP detection was
carried out by SSAHA-SNP. 863872 reads were annotated as STSs and
485941 SNPs were
annotated with alleles from the boxer and the breed or canid from
which the particular
read came. The validation rate for these SNPs was estimated at
approximately 98%.

A second set of SNPs was generated using a similar methodology
WGA-discovery (WGA) of Boxer/Poodle SNPs:
from the 1.5x poodle assembly (Kirkness 2003) were used instead of
except that the contigs
WGS reads. Since this
sequence lacked base quality scores, arbitrary quality scores of
phred 40 were assigned
before the poodle sequence was placed uniquely on the CanFam1.0
boxer assembly and SNP
detection was carried out by SSAHA-SNP. 1637780 SNPs were annotated
with alleles from the
boxer and the poodle. The validation rate for these SNPs was
estimated at approximately TBD%.

Internal-WGA-discovery (I-WGA):
A third set of SNPs were discovered by comparing reads in the WGA
assembly. SNPs were
defined as mismatch positions that had a base quality of >= 30 on
both reads in a region
that aligned without gaps, and with at most one additional mismatch
in the ten flanking
bases. For each allele, at least one additional read had to confirm
it. 731476 SNPs were
annotated with alleles between the two boxer alleles. The
validation rate for these SNPs
was estimated at approximately TBD%.

source
1..588
/organism="Canis familiaris"
/mol_type="Genomic DNA"
/strain="Rottweiler"
/db_xref="taxon:9615"
/map="1 22-531 17237500-17236987"
/clone.lib="Rottweiler"
<1..>588

STS
ORIGIN
Query Match 0.7%; Score 21; DB 7; Length 588;

Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3100 AAGAATTACTAAAAA 3120
|||||
Db 71 AAGAATTACTAAAAA 51
|||||

RESULT 146
LOCUS CQ431300/c 802 bp DNA linear PAT 28-JAN-2004
DEFINITION Sequence 16334 from Patent WO0151628.
ACCESSION CQ431300
VERSION CQ431300.1 GI:41383529
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.
TITLE Novel genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of breast cancer
JOURNAL Patent: WO 0151628-A 16334 19-JUL-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1..802
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 802;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2422 GGCTGGTCTTATGGAGGATAC 2442
|||||
Db 326 GGCTGGTCTTATGGAGGATAC 306
|||||

RESULT 147
LOCUS CQ422442/c 855 bp DNA linear PAT 28-JAN-2004
DEFINITION Sequence 7476 from Patent WO0151628.
ACCESSION CQ422442
VERSION CQ422442.1 GI:41374671
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.
TITLE Novel genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of breast cancer
JOURNAL Patent: WO 0151628-A 7476 19-JUL-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1..855
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 0.7%; Score 21; DB 2; Length 855;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2422 GGCTGGTCTTATGGAGGATAC 2442
|||||

```

Db      346  GCCTGGTCTCTATGGAGGATAC 326

RESULT 148
AF007822
LOCUS   Homo sapiens cytoplasmic Seprase truncated isoform mRNA, complete cds.
DEFINITION
ACCESSION AF007822
VERSION   1
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens

REFERENCE
  1 (bases 1 to 1143)
AUTHORS   Goldstein,L.A. and Chen,W.T.
TITLE     Identification of an alternatively spliced seprase mRNA that encodes a novel intracellular isoform
JOURNAL   J. Biol. Chem. 275 (4), 2554-2559 (2000)
PUBMED    10644713
AUTHORS   2 (bases 1 to 1143)
TITLE     Goldstein,L.A. and Chen,W.-T.
JOURNAL   Direct Submission
SUBMITTED (11-JUN-1997) Lombardi Cancer Center, Georgetown University Medical Center, 3970 Reservoir Road NW, Washington, DC 20007-2187, USA

FEATURES             Location/Qualifiers
     source           1..1143
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /cell_lines="LOX"
                     /cell_types="amelanotic melanoma"
                     399..1108
                     /note="fibroblast activation protein; intracellular isoform equivalent to carboxyl-terminal region of serine integral membrane protease Seprase; alternative splice variant"
                     /codon_start=1
                     /product="cytoplasmic Seprase truncated isoform"
                     /protein_id="AAFA3160.1"
                     /db_xref="GI:6649855"
                     /translation="MILFFQPDRSKKYPLLIQVYGPCSQSVRSVPANWISYLASKE
                     GMVIALVDGRTAQGDKLIYAVTRKLGVEYVEDQITAVRKFIEMGFIDEKRIATWGW
                     SYGVVSLALASGIGLFGKCIAGVAPVSSWEYIYASVYTERFNGLPTKDNLEHYKNST
                     VMARAEVRNYDILLIHGTADDNVHFQNSAQIAKALVNAQVDFOAMWYSDQNHGLSLG
                     STNHLVTHMTHFLKQCFSLSD"

ORIGIN
Query Match      0.7%; Score 21; DB 5; Length 1143;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2422  GCCTGGTCTCTATGGAGGATAC 2442
Db      689  GCCTGGTCTCTATGGAGGATAC 709

RESULT 149
HSM803473
LOCUS   Homo sapiens mRNA; cDNA DKF2p686G13158 (from clone DKF2p686G13158).
DEFINITION
ACCESSION AL832166
VERSION   1
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens

REFERENCE
  1 (bases 1 to 1615)
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
TITLE     1 (bases 1 to 1615)

FEATURES             Location/Qualifiers
     source           1..1615
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"

ORIGIN
Query Match      0.7%; Score 21; DB 2; Length 2255;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2422  GCCTGGTCTCTATGGAGGATAC 2442
Db      1771  GCCTGGTCTCTATGGAGGATAC 1791

Search completed: June 22, 2006, 08:46:52
Job time : 17308 secs

AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
TITLE     Direct Submission
JOURNAL   Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherrberg, GERMANY
COMMENT   Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKF2p686G13158) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source           1..1615
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clones="DKF2p686G13158"
                     /tissue_types="human colon endothel primary cell culture"
                     /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
                     DH10B; sites SfiIA + SfiIB"
                     /dev_stage="adult"
                     1513..1518
                     polyA_signal
                     polyA_site
                     ORIGIN
Query Match      0.7%; Score 21; DB 5; Length 1615;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2422  GCCTGGTCTCTATGGAGGATAC 2442
Db      823  GCCTGGTCTCTATGGAGGATAC 843

RESULT 150
CS141101
LOCUS   Sequence 3 from Patent WO2005071073.
DEFINITION
ACCESSION CS141101
VERSION   CS141101.1
KEYWORDS  GI:73531315
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS   McLean,P.A.
TITLE     Pap compositions and the use thereof for immunomodulation
JOURNAL   Patent WO 2005071073-A 3 04-AUG-2005;
Point Therapeutics, Inc. (US)
FEATURES             Location/Qualifiers
     source           1..2255
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

ORIGIN
Query Match      0.7%; Score 21; DB 2; Length 2255;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2422  GCCTGGTCTCTATGGAGGATAC 2442
Db      1771  GCCTGGTCTCTATGGAGGATAC 1791

Search completed: June 22, 2006, 08:46:52
Job time : 17308 secs

```

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☒ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)